



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2019

---

## On the dust temperatures of high-redshift galaxies

Liang, Lichen ; Feldmann, Robert ; Kereš, Dušan ; Scoville, Nick Z ; Hayward, Christopher C ;  
Faucher-Giguère, Claude-André ; Schreiber, Corentin ; Ma, Xiangcheng ; Hopkins, Philip F ; Quataert,  
Eliot

**Abstract:** Dust temperature is an important property of the interstellar medium (ISM) of galaxies. It is required when converting (sub)millimetre broad-band flux to total infrared luminosity (LIR), and hence star formation rate, in high-redshift galaxies. However, different definitions of dust temperatures have been used in the literature, leading to different physical interpretations of how ISM conditions change with, e.g. redshift and star formation rate. In this paper, we analyse the dust temperatures of massive ( $M_{\text{star}} > 10^{10} M_{\odot}$ )  $z = 2-6$  galaxies with the help of high-resolution cosmological simulations from the Feedback in Realistic Environments (fire) project. At  $z \sim 2$ , our simulations successfully predict dust temperatures in good agreement with observations. We find that dust temperatures based on the peak emission wavelength increase with redshift, in line with the higher star formation activity at higher redshift, and are strongly correlated with the specific star formation rate. In contrast, the mass-weighted dust temperature, which is required to accurately estimate the total dust mass, does not strongly evolve with redshift over  $z = 2-6$  at fixed IR luminosity but is tightly correlated with LIR at fixed  $z$ . We also analyse an ‘equivalent’ dust temperature for converting (sub)millimetre flux density to total IR luminosity, and provide a fitting formula as a function of redshift and dust-to-metal ratio. We find that galaxies of higher equivalent (or higher peak) dust temperature (‘warmer dust’) do not necessarily have higher mass-weighted temperatures. A ‘two-phase’ picture for interstellar dust can explain the different scaling relations of the various dust temperatures.

DOI: <https://doi.org/10.1093/mnras/stz2134>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-182559>

Journal Article

Published Version

Originally published at:

Liang, Lichen; Feldmann, Robert; Kereš, Dušan; Scoville, Nick Z; Hayward, Christopher C; Faucher-Giguère, Claude-André; Schreiber, Corentin; Ma, Xiangcheng; Hopkins, Philip F; Quataert, Eliot (2019). On the dust temperatures of high-redshift galaxies. *Monthly Notices of the Royal Astronomical Society*, 489(1):1397-1422.

DOI: <https://doi.org/10.1093/mnras/stz2134>

# On the dust temperatures of high-redshift galaxies

Lichen Liang<sup>1</sup>,<sup>1</sup>★ Robert Feldmann<sup>1</sup>, Dušan Kereš<sup>2</sup>, Nick Z. Scoville,<sup>3</sup>  
 Christopher C. Hayward,<sup>4</sup> Claude-André Faucher-Giguère,<sup>5</sup> Corentin Schreiber,<sup>6,7</sup>  
 Xiangcheng Ma<sup>3,8</sup>, Philip F. Hopkins<sup>3</sup> and Eliot Quataert<sup>8</sup>

<sup>1</sup>*Institute for Computational Science, University of Zurich, Zurich CH-8057, Switzerland*

<sup>2</sup>*Department of Physics, Centre for Astrophysics and Space Sciences, University of California at San Diego, La Jolla, CA 92093, USA*

<sup>3</sup>*Cahill Center for Astrophysics, California Institute of Technology, 1216 East California Boulevard, Pasadena, CA 91125, USA*

<sup>4</sup>*Centre for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA*

<sup>5</sup>*Department of Physics and Astronomy and CIERA, Northwestern University, Evanston, IL 60208, USA*

<sup>6</sup>*Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*

<sup>7</sup>*Leiden Observatory, Leiden University, NL-2300 RA Leiden, the Netherlands*

<sup>8</sup>*Department of Astronomy, 501 Campbell Hall, University of California, Berkeley, CA 94720, USA*

Accepted 2019 July 30. Received 2019 July 17; in original form 2019 March 12

## ABSTRACT

Dust temperature is an important property of the interstellar medium (ISM) of galaxies. It is required when converting (sub)millimetre broad-band flux to total infrared luminosity ( $L_{\text{IR}}$ ), and hence star formation rate, in high-redshift galaxies. However, different definitions of dust temperatures have been used in the literature, leading to different physical interpretations of how ISM conditions change with, e.g. redshift and star formation rate. In this paper, we analyse the dust temperatures of massive ( $M_{\text{star}} > 10^{10} M_{\odot}$ )  $z = 2$ – $6$  galaxies with the help of high-resolution cosmological simulations from the *Feedback in Realistic Environments* (FIRE) project. At  $z \sim 2$ , our simulations successfully predict dust temperatures in good agreement with observations. We find that dust temperatures based on the peak emission wavelength increase with redshift, in line with the higher star formation activity at higher redshift, and are strongly correlated with the specific star formation rate. In contrast, the *mass-weighted* dust temperature, which is required to accurately estimate the total dust mass, does not strongly evolve with redshift over  $z = 2$ – $6$  at fixed IR luminosity but is tightly correlated with  $L_{\text{IR}}$  at fixed  $z$ . We also analyse an ‘equivalent’ dust temperature for converting (sub)millimetre flux density to total IR luminosity, and provide a fitting formula as a function of redshift and dust-to-metal ratio. We find that galaxies of higher equivalent (or higher peak) dust temperature (‘warmer dust’) do not necessarily have higher mass-weighted temperatures. A ‘two-phase’ picture for interstellar dust can explain the different scaling relations of the various dust temperatures.

**Key words:** galaxies: evolution – galaxies: high-redshift – galaxies: ISM – submillimetre: galaxies.

## 1 INTRODUCTION

Astrophysical dust, originating from the condensation of metals in stellar ejecta, is pervasive in the interstellar medium (ISM) of galaxies in both local and distant Universe (e.g. Smail, Ivison & Blain 1997; Blain et al. 1999; Chapman et al. 2005; Capak et al. 2011; Riechers et al. 2013; Weiß et al. 2013; Capak et al. 2015; Watson et al. 2015; Ivison et al. 2016; Laporte et al. 2017; Venemans et al. 2017; Hashimoto et al. 2018; Pavesi et al. 2018; Zavala et al. 2018a; Izumi et al. 2019, and references therein).

Dust scatters and absorbs UV-to-optical light, and therefore strongly impacts the observed flux densities and the detectability of galaxies at these wavelengths (e.g. Kinney et al. 1993; Calzetti, Kinney & Storchi-Bergmann 1994; Calzetti et al. 2000; Kriek & Conroy 2013; Narayanan et al. 2018b). Despite that it accounts for no more than a few per cent of the total ISM mass (Draine et al. 2007), dust also plays a key role in star formation process of galaxies (e.g. Cazaux & Tielens 2002; Murray, Quataert & Thompson 2005; McKee & Ostriker 2007; Hopkins, Quataert & Murray 2012). Constraining and understanding dust properties of galaxies are therefore essential for proper interpretation of the multiwavelength data from observations and for facilitating our understanding of galaxy formation and evolution.

\* E-mail: lliang@physik.uzh.ch

Much of the stellar emission of star-forming galaxies is absorbed by dust grains and re-emitted at infrared (IR)-to-millimetre (mm) wavelengths as thermal radiation, encoding important information about dust and galactic properties, such as dust mass, total IR luminosity<sup>1</sup> ( $L_{\text{IR}}$ ), and star formation rate (SFR) (e.g. Chary & Elbaz 2001; Dale & Helou 2002; Siebenmorgen & Krügel 2007; da Cunha, Charlot & Elbaz 2008; Ivison et al. 2010; Walcher et al. 2011; Casey 2012; Magdis et al. 2012; Casey, Narayanan & Cooray 2014; Scoville et al. 2016; Schreiber et al. 2018). The advent of the new facilities in the past two decades, such as the *Spitzer Space Telescope* (Fazio et al. 2004), *Herschel Space Observatory* (Pilbratt et al. 2010), the Submillimetre Common-user Bolometer Array (SCUBA) camera on the James Clerk Maxwell Telescope (JCMT) (Holland et al. 1999, 2013), the AzTEC millimetre camera on the Large Millimeter Telescope (LMT) (Wilson et al. 2008), the South Pole Telescope (SPT) (Carlstrom et al. 2011), and the Atacama Large Millimeter/Sub-millimeter Array (ALMA) has triggered significant interests in the study of ISM dust. In particular, observations with the Photodetector Array Camera and Spectrometer (PACS; Poglitsch et al. 2010) and the Spectral and Photometric Imaging Receiver (SPIRE; Griffin et al. 2010) instruments aboard *Herschel* made it possible to study the 70–500  $\mu\text{m}$  wavelength range where most of the Universe’s obscured radiation emerges, and many dust-enshrouded, previously unreported objects at distant space have been uncovered through the wide-area extragalactic surveys (e.g. Eales et al. 2010; Lutz et al. 2011; Oliver et al. 2012). Far-infrared (FIR)-to-mm spectral energy distribution (SED) modelling of dust emission has therefore become possible for objects at high redshift ( $z \sim 4$ , Weiß et al. 2013; Ivison et al. 2016; Schreiber et al. 2018) and various dust properties can be extracted using SED fitting techniques (Walcher et al. 2011).

The Rayleigh–Jeans (RJ) side (i.e.  $\lambda > \lambda_{\text{peak}}$ , where  $\lambda_{\text{peak}}$  is the wavelength of the peak emission) of the dust SED can empirically be well fitted by a single-temperature ( $T$ ) modified blackbody (MBB) function (Hildebrand 1983). However, the shape of the Wien side (i.e.  $\lambda < \lambda_{\text{peak}}$ ) of the SED, which is tied to the warm dust component in vicinity of the young stars and active galactic nuclei (AGNs), has a much larger variety (e.g. Kirkpatrick et al. 2012; Symeonidis et al. 2013). Studies have shown that *one* single- $T$  MBB function cannot well fit *both* sides of the observed SEDs (Casey 2012). Motivated by this fact, fitting functions of multi- $T$  components have been adopted (e.g. Dunne & Eales 2001; Blain, Barnard & Chapman 2003; Kovács et al. 2010; Casey 2012; Dale et al. 2012; Galametz et al. 2012; Kirkpatrick et al. 2012; da Cunha et al. 2013, 2015). Meanwhile, empirical SED templates have been developed based on an assumed distribution of interstellar radiation intensity ( $U$ ) incident on dust grains (e.g. Dale et al. 2001; Draine & Li 2007; Galliano et al. 2011; Schreiber et al. 2018). Both approaches can produce functional shape of dust SED that better matches the observed photometry of galaxies compared to a single- $T$  MBB function.

At high redshift ( $z \gtrsim 2$ ), however, it is more common that a galaxy has only a few (two or three) reliable photometric data points in its dust continuum so that SED fitting by multi- $T$  functions or more sophisticated SED templates is not possible. Therefore, the widely adopted approach is to simply fit the available data points with *one* single- $T$  MBB function (e.g. Magnelli et al. 2012; Symeonidis et al.

2013; Magnelli et al. 2014; Simpson et al. 2017; Thomson et al. 2017). The  $T$  parameter that yields the best fit is then often referred to as the ‘dust temperature’ of the galaxy in the literature. We specify this definition of dust temperature as the ‘effective’ temperature ( $T_{\text{eff}}$ ) in this paper. Another temperature also often used is the ‘peak’ temperature ( $T_{\text{peak}}$ ), which is defined based on the emission peak of the best-fitting SED assuming Wien’s displacement law (Casey et al. 2014). These observationally derived dust temperatures (both  $T_{\text{eff}}$  and  $T_{\text{peak}}$ ) can depend on the assumed functional form of SED as well as the adopted photometry (Casey 2012; Casey et al. 2014). Despite that it is unclear how well these simplified fitting functions represent the true SED shape of high-redshift galaxies and the physical interpretation of the derived temperatures is not obvious, this approach is frequently used to analyse large statistical samples of data (e.g. Chapman et al. 2005; Hwang et al. 2010; Symeonidis et al. 2013; Swinbank et al. 2014; da Cunha et al. 2015; Casey et al. 2018a).

The scaling relations of  $T_{\text{peak}}$  ( $T_{\text{eff}}$ ) and other dust/galaxy properties, including the  $L_{\text{IR}}$ –temperature and specific star formation rate (sSFR)–temperature relations, may be related to the physical conditions of the star-forming regions in distant galaxies and have attracted much attention (e.g. Magdis et al. 2012; Magnelli et al. 2012, 2014; Lutz 2014; Casey et al. 2018b; Schreiber et al. 2018). While observational studies derive dust temperatures in a variety of ways (using different fitting functions and/or different photometry), they generally infer that the temperature increases with  $L_{\text{IR}}$  and sSFR of galaxies. Accurate interpretation of these observed scaling relationships requires a knowledge of how different dust and galaxy properties (e.g. stellar mass, SFR, and dust mass) shape the dust SED (Draine & Li 2007; Groves et al. 2008; Scoville 2013; Safarzadeh et al. 2016), and hence the derived dust temperatures. Radiative transfer (RT) analyses of galaxy models are important tools for understanding these temperatures since the intrinsic properties of the simulated galaxies are known (e.g. Narayanan et al. 2010; Hayward et al. 2011, 2012; Hayward & Smith 2015; Narayanan et al. 2015; Camps et al. 2016; Liang et al. 2018; Ma et al. 2019).

One important question is how the derived temperatures are related to the physical, *mass-weighted* dust temperature ( $T_{\text{mw}}$ ). Observations of the local galaxies have shown that the bulk of the ISM dust remains at low temperature (Dunne & Eales 2001; Harvey et al. 2013; Lombardi et al. 2014). The cold dust component determines  $T_{\text{mw}}$  of the galaxy, which sets the shape of the RJ tail. For distant galaxies, it is very challenging to measure  $T_{\text{mw}}$  due to the limit of resolution. However, a good estimate of  $T_{\text{mw}}$  is important for deriving the ISM masses of high-redshift galaxies via the RJ method (e.g. Scoville et al. 2014, 2016, 2017b). It is unclear whether, or how, one can infer  $T_{\text{mw}}$  from the observationally derived temperatures. Alternatively, one can simply adopt a constant value if  $T_{\text{mw}}$  has relatively small variation among different galaxies, given that the mass estimates scale only *linearly* with  $T_{\text{mw}}$  (Scoville et al. 2016). If that is the case, it can also be one major advantage of the RJ approach because the main difficulty of the traditional CO method is the large uncertainty of the CO-to- $\text{H}_2$  conversion factor (e.g. Shetty et al. 2011; Feldmann, Gnedin & Kravtsov 2012; Carilli & Walter 2013). RT analyses are useful for understanding the relation between the derived temperatures and  $T_{\text{mw}}$  (Liang et al. 2018).

Over the past two decades, many ground-based galaxy surveys at (sub)mm wavelengths (e.g. SCUBA, AzTEC, SPT, and ALMA) that are complementary to *Herschel* observations (e.g. Smail et al.

<sup>1</sup>In this paper,  $L_{\text{IR}}$  is defined as the luminosity density integrated over the 8–1000  $\mu\text{m}$  wavelength interval.

1997; Dunne et al. 2000; Geach et al. 2013; Karim et al. 2013; Swinbank et al. 2014; Aravena et al. 2016; Bouwens et al. 2016; Hatsukade et al. 2016; Walter et al. 2016; Dunlop et al. 2017; Geach et al. 2017; Franco et al. 2018, and references therein) have been carried out. Deep (sub)mm surveys are capable of probing less actively star-forming (SFRs  $\lesssim 100 M_{\odot} \text{ yr}^{-1}$ ) galaxies at  $z \lesssim 4$  (e.g. Hatsukade et al. 2013; Chen et al. 2014; Ono et al. 2014; Zavala et al. 2018b). Furthermore, they are effective at uncovering sources at  $z > 4$  thanks to the effect of ‘negative  $K$ -correction’ (e.g. Capak et al. 2015; Carniani et al. 2015; Fujimoto et al. 2016; Laporte et al. 2017; Casey et al. 2018b). The (sub)mm-detected sources do not necessarily have *Herschel* counterparts. Deriving  $L_{\text{IR}}$  (and hence SFR) of these sources from a single (sub)mm flux density ( $S$ ) requires adopting an assumed dust temperature, which we refer to as ‘equivalent’ temperature ( $T_{\text{eqv}}$ ) in this paper, along with an assumed (simplified) functional shape of the dust SED (Bouwens et al. 2016; Casey et al. 2018b).  $T_{\text{eqv}}$  is conceptually different from  $T_{\text{eff}}$  introduced above because the former is an *assumed* quantity for extrapolating  $L_{\text{IR}}$  from a *single* data point while the latter is a *derived* quantity through SED fitting to *multiple* data points.

A good estimate of  $T_{\text{eqv}}$  is important for translating the information (e.g. source number counts) extracted from the ALMA blind surveys to the *obscured* cosmic star formation density at  $z \gtrsim 4$  (Casey et al. 2018a,b; Zavala et al. 2018c), where currently only reliable data from rest-frame UV measurements are available (Madau & Dickinson 2014). One common finding of the recent (sub)mm blind surveys is a dearth of faint sources at these early epochs, as noted by Casey et al. (2018a) (see also the references therein). This can suggest that the early Universe is relatively dust-poor and only a small fraction of stellar emission is absorbed and re-emitted by dust (Casey et al. 2018a). Alternatively, it could also be accounted for by a significantly higher  $T_{\text{eqv}}$  at high redshifts, meaning that galaxies of the same  $L_{\text{IR}}$  appear to be fainter in the (sub)mm bands (cf. Capak et al. 2015; Bouwens et al. 2016; Faisst et al. 2017; Fudamoto et al. 2017; Narayanan et al. 2018a). Hence, understanding how  $T_{\text{eqv}}$  evolves with redshift and how it depends on different galaxy properties are crucial for constraining the total amount of dust and the amount of *obscured* star formation density in the early Universe (Casey et al. 2018a,b).

In this paper, we study in detail the observational and the physical (mass-weighted) dust temperatures with the aid of high-resolution cosmological galaxy simulations. In particular, we study a sample of massive ( $M_{\text{star}} > 10^{10} M_{\odot}$ )  $z = 2$ – $6$  galaxies from the FIRE project<sup>2</sup> (Hopkins et al. 2014) with dust RT modelling. This sample contains galaxies with  $L_{\text{IR}}$  ranging over two orders of magnitude, from  $10^{10}$  to  $10^{12} L_{\odot}$  and few dust-rich, ultra-luminous ( $L_{\text{IR}} \gtrsim 10^{12} L_{\odot}$ ) galaxies at  $z \sim 2$  that are candidates for both *Herschel*- and submm-detected objects. A lot of them have  $L_{\text{IR}} \sim \text{a few} \times 10^{11} L_{\odot}$ , which is accessible by *Herschel* using stacking techniques (e.g. Thomson et al. 2017; Schreiber et al. 2018). Our sample also contains fainter galaxies at  $z = 2$ – $6$  with observed flux densities  $S_{870 \mu\text{m}} (S_{1.2\text{mm}}) \gtrsim 0.1 \text{ mJy}$ , which could be potentially detected with ALMA. We calculate and explicitly compare their  $T_{\text{mw}}$  with the observationally derived temperatures ( $T_{\text{peak}}$  or  $T_{\text{eff}}$ ), as well as their scaling relationships with several galaxy properties. We also provide the prediction for  $T_{\text{eqv}}$  that is needed for deriving  $L_{\text{IR}}$  of galaxy from its observed single-band (sub)mm flux.

The paper is structured as follows. In Section 2, we introduce the simulation details and the methodology of RT modelling. In

Section 3, we provide the various definitions of dust temperature in detail, discuss the impact of dust temperature on SED shape, and compare the specific predictions of our simulations with observations. In Section 4, we focus on the conversion from single-band (sub)mm broad-band flux to  $L_{\text{IR}}$  and provide useful fitting formulae. In Section 5, we discuss the observational implications of our findings. We summarize and conclude in Section 6.

Throughout this paper, we adopt cosmological parameters in agreement with the 9-yr data from the Wilkinson Microwave Anisotropy Probe (Hinshaw et al. 2013), specifically  $\Omega_{\text{m}} = 0.2821$ ,  $\Omega_{\Lambda} = 0.7179$ , and  $H_0 = 69.7 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

## 2 SIMULATION METHODOLOGY

In this section, we introduce our simulation methodology. In Section 2.1, we briefly summarize the details of the cosmological hydrodynamic simulations from which our galaxy sample is extracted. In Section 2.2, we introduce the methodology of our dust RT analysis and present mock images produced with SKIRT.

### 2.1 Simulation suite and sample

We extract our galaxy sample from the MASSIVEFIRE cosmological zoom-in suite (Feldmann et al. 2016; Feldmann et al. 2017), which is part of the Feedback in Realistic Environments (FIRE) project.

The initial conditions for the MASSIVEFIRE suites are generated using the MUSIC code (Hahn & Abel 2011) within a  $(100 \text{ Mpc/h})^3$  comoving periodic box with the WMAP cosmology. From a low-resolution (LR) dark matter (DM)-only run, isolated haloes of a variety of halo masses, accretion history, and environmental overdensities are selected. Initial conditions for the ‘zoom-in’ runs use a convex hull surrounding all particles within  $3R_{\text{vir}}$  at  $z = 2$  of the chosen halo defining the Lagrangian high-resolution (HR) region. The mass resolution of the default HR runs is  $m_{\text{DM}} = 1.7 \times 10^5 M_{\odot}$  and  $m_{\text{gas}} = 3.3 \times 10^4 M_{\odot}$ , respectively. The initial mass of the star particle is set to be the same as the parent gas particle from which it is spawned in the simulations.

The simulations are run with the gravity-hydrodynamics code GIZMO<sup>3</sup> (FIRE-1 version) in the Pressure–energy Smoothed Particle Hydrodynamics (‘P-SPH’) mode (Hopkins 2015), which improves the treatment of fluid mixing instabilities and includes various other improvements to the artificial viscosity, artificial conductivity, higher order kernels, and time-stepping algorithm designed to reduce the most significant known discrepancies between SPH and grid methods (Hopkins 2013). Gas that is locally self-gravitating and has density over  $5 \text{ cm}^{-3}$  is assigned an SFR  $\dot{\rho} = f_{\text{mol}} \rho / t_{\text{ff}}$ , where  $f_{\text{mol}}$  is the self-shielding molecular mass fraction. The simulations explicitly incorporate several different stellar feedback channels (but not feedback from supermassive black holes) including (1) local and long-range momentum flux from radiative pressure, (2) energy, momentum, mass, and metal injection from supernovae (Types Ia and II), (3) and stellar mass-loss (both OB and AGB stars), and (4) photoionization and photoelectric heating processes. We refer the reader to Hopkins et al. (2014) for details.

In this study, we analyse 18 massive ( $10^{10} < M_{\text{star}} < 10^{11.3} M_{\odot}$  at  $z = 2$ ) central galaxies (from Series A, B, and C in Feldmann et al. 2017) and their most massive progenitors (MMPs) up to  $z = 6$ , identified using the Amiga Halo Finder (Gill, Knebe & Gibson

<sup>2</sup>fire.northwestern.edu

<sup>3</sup>A public version of GIZMO is available at <http://www.tapir.caltech.edu/phopkins/Site/GIZMO.html>.



2004; Knollmann & Knebe 2009). These galaxies were extracted from the haloes selected from the LR DM-only run. In order to better probe the dusty, IR-luminous galaxies at the extremely high-redshift ( $z > 4$ ) Universe, we also include another 11 massive ( $10^{10} < M_{\text{star}} < 10^{11} M_{\odot}$  at  $z = 6$ ) galaxies extracted from a different set of MASSIVEFIRE simulations that stop at  $z = 6$ , which are presented here for the first time. The latter were run with the same physics, initial conditions, numerics, and spatial and mass resolution, but were extracted from larger simulation boxes (400 and 762 Mpc  $h^{-1}$  on a side, respectively).

FIRE simulations successfully reproduce a variety of observed galaxy properties relevant for this work, such as the stellar-to-halo-mass relation (Hopkins et al. 2014; Feldmann et al. 2017), the sSFRs of galaxies at the cosmic noon ( $z \sim 2$ ) (Hopkins et al. 2014; Feldmann et al. 2016), the stellar mass–metallicity relation (Ma et al. 2016), and the submm flux densities at 850  $\mu\text{m}$  (Liang et al. 2018).

## 2.2 Predicting dust SED with SKIRT

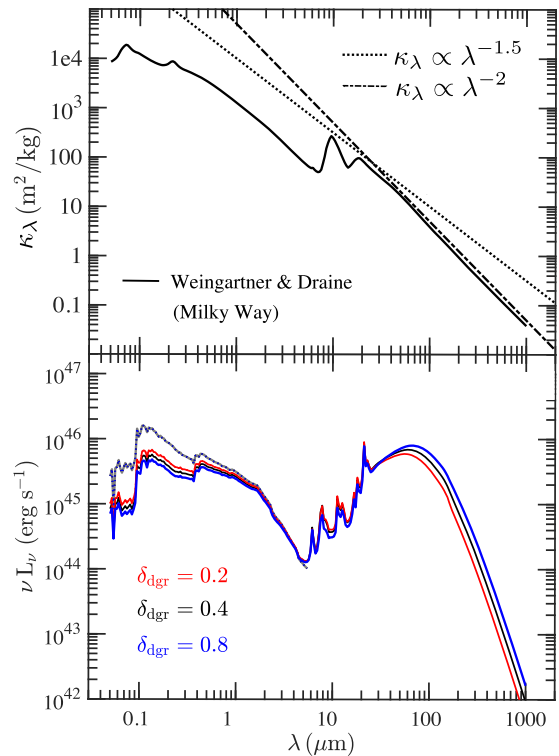
We generate the UV-to-mm SED using the open source<sup>4</sup> 3D dust Monte Carlo RT code SKIRT (Baes et al. 2011; Baes & Camps 2015). SKIRT accounts for absorption and anisotropic scattering of dust and self-consistently calculates the dust temperature. We follow the approach by Camps et al. (2016) (see also Trayford et al. 2017) to prepare our galaxy snapshots as RT input models.

Each star particle in the simulation is treated as a ‘single stellar population’ (SSP). The spectrum of a star particle in the simulation is assigned using STARBURST99 SED libraries. In our default RT model, every star particle is assigned an SED according to the age and metallicity of the particle.

While our simulations have better resolution than many previous simulations modelling infrared and submm emission (e.g. Narayanan et al. 2010; Hayward et al. 2011; De Looze et al. 2014) and can directly incorporate various important stellar feedback processes, they are still unable to resolve the emission from H II and photodissociation regions (PDR) from some of the more compact birth-clouds surrounding star-forming cores. The time-average spatial scale of these H II+PDR regions typically varies from  $\sim 5$  to  $\sim 800$  pc depending on the local physical conditions (Jonsson, Groves & Cox 2010). Hence, in our alternative RT model, star particles are split into two sets based on their age. Star particles that formed less than 10 Myr ago are identified as ‘young star-forming’ particles, while older star particles are treated as above. We follow Camps et al. (2016) in assigning a source SED from the MAPPINGSIII (Groves et al. 2008) family to young star-forming particles to account for the pre-processing of radiation by birth-clouds. Dust associated with the birth-clouds is removed from the neighbouring gas particles to avoid double-counting (see also Camps et al. 2016).

We present in Sections 3 and 4 the results from our default (‘no birth-cloud’) model. In Section 5, we will show that none of our results are qualitatively altered if we adopt the alternative RT model and account for unresolved birth-clouds.

Our RT analysis uses  $10^6$  photon packets for each stage. We use an octree for the dust grid and keep subdividing grid cells until the cell contains less than  $f = 3 \times 10^{-6}$  of the total dust mass and the V-band optical depth in each cell is less than unity. The highest grid level corresponds to a cell width of  $\sim 20$  pc, i.e. about



**Figure 1.** *Upper panel:* Dust opacity curve of the Weingartner & Draine (2001) dust model that is used for analysis in this work (solid line). The dashed and dash-dotted lines show the asymptotic power law  $\kappa_{\lambda} \propto \lambda^{-1.5}$  and  $\kappa_{\lambda} \propto \lambda^{-2}$ , respectively. *Lower panel:* SEDs of a selected  $z = 2$  MASSIVEFIRE galaxy. The red, black, and blue curves show results for  $\delta_{\text{dgr}} = 0.2$ , 0.4, and 0.8, respectively. The grey curve shows the intrinsic stellar emission. About half of the stellar radiative energy of this galaxy is absorbed and re-emits at IR.

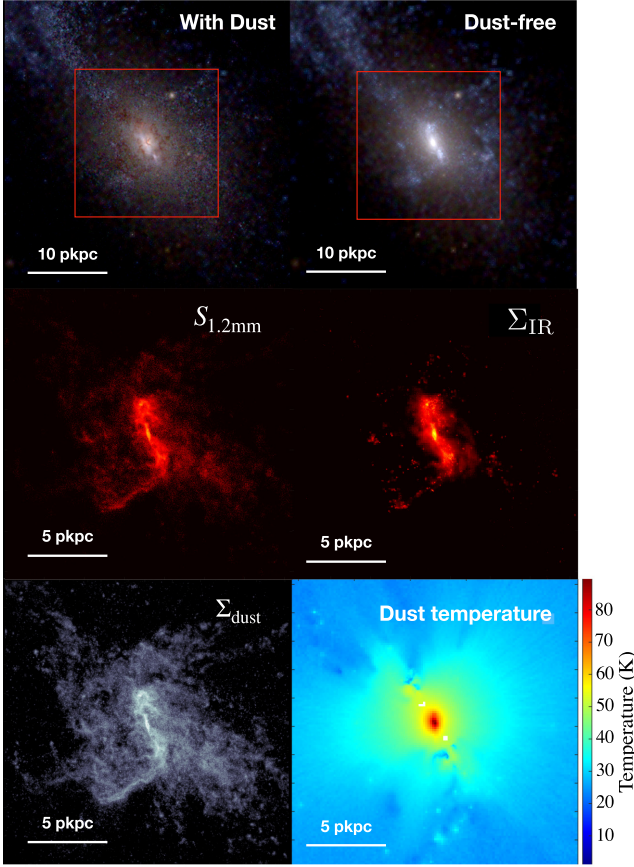
twice the minimal SPH smoothing length. For all the analysis in this paper, we adopt the Weingartner & Draine (2001) dust model with Milky Way size distribution for the case of  $R_V = 3.1$ . At FIR, the dust opacity can be well described by a power law,  $\kappa_{\lambda} \propto 0.05 (\lambda/870 \mu\text{m})^{-\beta} \text{ m}^2 \text{ kg}^{-1}$ , where  $\beta \approx 2.0$  (see the upper panel of Fig. 1) is the dust emissivity spectral index (consistent with the observational constraints, e.g. Dunne et al. 2000; Draine et al. 2007). Gas hotter than  $10^6$  K is assumed to be dust-free due to sputtering (Hirashita et al. 2015). We self-consistently calculate the self-absorption of dust emission and include the transient heating function to calculate non-local thermal equilibrium (NLTE) dust emission by transiently heated small grains and PAH molecules (Baes et al. 2011). Transient heating influences the rest-frame mid-infrared (MIR) emission ( $\lesssim 80 \mu\text{m}$ ) but has minor impact on the FIR and (sub)mm emission (Behrens et al. 2018). SKIRT outputs  $T_{\text{mw}}$  for each cell that is obtained by averaging the temperature over grains of different species (composition and size). A galaxy-wide dust temperature is calculated by mass-weighting  $T_{\text{mw}}$  of each cell in the galaxies. At high redshift ( $z > 4$ ), the radiation field from the cosmic microwave background (CMB) starts to affect the temperature of the cold ISM. We account for the CMB by adopting a corrected dust temperature (da Cunha et al. 2013)

$$T_{\text{dust}}^{\text{corr}}(z) = [T_{\text{dust}}^{4+\beta} + T_{\text{CMB}}^{4+\beta}(z) - T_{\text{CMB}}^{4+\beta}(z=0)]^{1/(4+\beta)}, \quad (1)$$

where  $T_{\text{CMB}}(z) = 2.73(1+z)$  K is the CMB temperature at  $z$ .

For this study, we assume that dust mass traces metal mass in the ISM, and adopt a constant dust-to-metal mass ratio  $\delta_{\text{dgr}} = 0.4$

<sup>4</sup>SKIRT code repository: <https://github.com/skirt>.



**Figure 2.** Example of the RT analysis applied to a  $z = 2$  MASSIVEFIRE galaxy. *Upper panels:* UVJ image with (left) and without (right) the effect of dust extinction. *Middle panels:* Normalized  $S_{1.2\text{mm}}$  (left) and normalized  $\Sigma_{\text{IR}}$  (right). Compared with  $S_{1.2\text{mm}}$ ,  $\Sigma_{\text{IR}}$  traces more tightly to the star-forming regions. *Lower panels:* Dust surface density (left) and dust temperature weighted along the line of sight, weighted by mass (right). The middle and lower panels show the result for the zoomed-in region enclosed by the red box in the upper panels.

(Dwek 1998; Draine et al. 2007; Watson 2011) for our *fiducial* analysis. We also try two different cases where  $\delta_{\text{dtr}} = 0.2$  and  $0.8$ , and throughout the paper, we refer to these two *dust-poor* and *dust-rich* cases, respectively. In the lower panel of Fig. 1, we show the galaxy SED for the three models.  $L_{\text{IR}}$  increases when  $\delta_{\text{dtr}}$  increases because a higher optical depth leads to more absorption of stellar light and more re-emission at IR.

SKIRT produces spatially resolved, multiwavelength rest-frame SEDs for each galaxy snapshot observed from multiple viewing angles. For the analysis in this paper, SEDs are calculated on an equally spaced logarithmic wavelength grid ranging from rest frame  $0.005$  to  $1000 \mu\text{m}$ . We convolve the simulated SED output from SKIRT with the transmission functions of the PACS (70, 100, 160  $\mu\text{m}$ ), SPIRE (250, 350, 500  $\mu\text{m}$ ), SCUBA-2 (450, 850  $\mu\text{m}$ ), and ALMA band 6 (870  $\mu\text{m}$ ) and 7 (1.2 mm) to yield the broad-band flux density for each band.

We show in Fig. 2 the result of running SKIRT on one of our galaxies. In particular, we show a composite  $U, V, J$  false-colour image with and without accounting for dust absorption, scattering, and emission. We also show the image of ALMA 1.2 mm flux density, total IR luminosity, dust surface density, and temperature. It can be seen that the 1.2 mm flux density traces the dust mass

distribution, while IR luminosity appears to be more localized to the high-temperature region, since it is expected to be sensitive to temperature ( $L \sim MT^{4+\beta}$ ). The local intensity of radiation, the dust temperatures, and the dust density all peak in the central region of the galaxy.

### 3 UNDERSTANDING DUST TEMPERATURE AND ITS SCALING RELATIONS

In this section, we at first review the different ways of defining galaxy dust temperature that have been used in different observational and theoretical studies (Section 3.1), and compute the different temperatures for the MASSIVEFIRE sample (Section 3.2). We compare the calculated dust temperature(s) of the simulated galaxies with recent observational data (Section 3.3). Finally, we reproduce several observed scaling relations (e.g.  $L_{\text{IR}}$  versus temperature, sSFR versus temperature) with the simulated galaxies and provide physical insights for these relations (Section 3.4).

#### 3.1 Defining dust temperature

Dust temperature has been defined in different ways by observational and theoretical studies. Here, we focus on four different possibilities, which we call *mass-weighted*, *peak*, *effective*, and *equivalent* dust temperature.

##### 3.1.1 Mass-weighted dust temperature $T_{\text{mw}}$

$T_{\text{mw}}$  is the physical, mass-weighted temperature of dust in the ISM.  $T_{\text{mw}}$  is often explicitly discussed in theoretical studies where dust RT modelling is applied to the snapshots from the galaxy simulations, and dust temperature is calculated using LTE (for large grains) and non-LTE (for small grains and PAH molecules) approaches (e.g. Behrens et al. 2018; Liang et al. 2018).

##### 3.1.2 Peak dust temperature $T_{\text{peak}}$

The peak dust temperature is defined based on the wavelength  $\lambda_{\text{peak}}$  at which the FIR spectral flux density reaches a maximum (e.g. Casey et al. 2014)

$$T_{\text{peak}} = \frac{2.90 \times 10^3 \mu\text{m K}}{\lambda_{\text{peak}}} \quad (2)$$

The peak wavelength  $\lambda_{\text{peak}}$  is commonly derived from fitting the SED to a specific functional form, for instance, an MBB, see below.  $\lambda_{\text{peak}}$  (and  $T_{\text{peak}}$ ) in practice depends on the adopted functional form as well as the broad-band photometry used in the fit (Casey 2012; Casey et al. 2014).

##### 3.1.3 Effective dust temperature $T_{\text{eff}}$

The effective temperature is obtained by fitting the SED with a parametrized function. The effective temperature is thus a fit parameter, and like  $T_{\text{peak}}$ , depends on both the adopted functional form and the broad-band photometry.

For most observed SEDs, the RJ side of the dust continuum can be well described by a generalized modified-blackbody function (G-MBB) of the form (Hildebrand 1983)

$$S_{\nu_0}(T) = A \frac{(1+z)}{d_L^2} (1 - e^{-\tau_{\nu}}) B_{\nu}(T) \quad (3)$$

$$= \frac{1 - e^{-\tau_\nu}}{\tau_\nu} \frac{(1+z)}{d_L^2} \kappa_\nu M_{\text{dust}} B_\nu(T), \quad (4)$$

where  $\nu_o$  is the observer's frequency,  $\nu = \nu_o(1+z)$  is the rest-frame frequency,  $\tau_\nu$  is the dust optical depth at  $\nu$ ,<sup>5</sup>  $\kappa_\nu$  is the dust opacity (per unit dust mass) at  $\nu$ ,  $B_\nu(T)$  is the Planck function,  $A$  is the surface area of the emitting source, and  $d_L$  is the luminosity distance from the source.  $\tau_\nu$  is often fitted by a power law at FIR wavelengths, i.e.  $\tau_\nu = (\nu/\nu_1)^\beta$ , where  $\beta$  is the spectral emissivity index and  $\nu_1$  is the frequency where optical depth is unity. Observational evidence has shown that the value of  $\nu_1$  can differ between galaxies (Gonzalez-Alfonso et al. 2004; Scoville et al. 2017a; Simpson et al. 2017). In principle,  $\nu_1$  can be determined from SED fitting given full FIR-to-mm coverage (Casey 2012). However, in practice, it is often taken to be a constant,  $\sim 1.5\text{--}3$  THz (i.e.  $\lambda_1 = c/\nu_1 = 100\text{--}200$   $\mu\text{m}$ ) (e.g. Draine 2006; Conley et al. 2011; Riechers et al. 2013; Symeonidis et al. 2013; Casey et al. 2014; Casey et al. 2018a,b; Zavala et al. 2018b).

The Wien side of the dust emission is expected to be strongly affected by the warm dust component in the vicinity of the star-forming regions, which can significantly boost the luminosity of galaxy with only a small mass fraction (e.g. Dunne & Eales 2001; Harvey et al. 2013), knowing  $L \sim MT^{4+\beta}$ . Observations also show a variety of SED shapes at MIR (e.g. Kirkpatrick et al. 2012; Symeonidis et al. 2013). To better account for the emission at MIR, Casey (2012) introduced a simple (truncated) power-law component to equation (3), giving rise to a G-MBB with an additional power-law component (GP-MBB)

$$S_{\nu_0}(T) = A \frac{(1+z)}{d_L^2} \left[ (1 - e^{-\tau_\nu}) B_\nu(T) + N_{\text{pl}} \nu^{-\alpha} e^{-(\nu_c(T)/\nu)^2} \right], \quad (5)$$

where  $N_{\text{pl}}$  is the normalization factor,  $\alpha$  is the power-law index, and  $\nu_c$  is a cut-off frequency where the power-law term turns over and no longer dominates the emission at MIR. We allow  $N_{\text{pl}}$  as a free parameter, fix  $\alpha = 2.5$ , and adopt the functional form of  $\nu_c(T)$  provided by Casey (2012). The latter were constrained by fitting the observational data of a sample of local IR-luminous galaxies from the Great-Origins All Sky LIRG Survey (GOALS; Armus et al. 2009).

In the optically thin regime ( $\tau \ll 1$ ), equation (5) reduces to the optically thin modified blackbody function (OT-MBB) (see e.g. Hayward et al. 2011)

$$\begin{aligned} S_{\nu_0} &= \frac{(1+z)}{d_L^2} \kappa_\nu M_{\text{dust}} B_\nu(T) \\ &= \frac{(1+z)}{d_L^2} \kappa_{870} \left( \frac{\nu}{\nu_{870}} \right)^\beta M_{\text{dust}} B_\nu(T) \\ &= C_\nu(z) M_{\text{dust}} B_\nu(T), \end{aligned} \quad (6)$$

where  $\kappa_{870}$  is the opacity at 870  $\mu\text{m}$  ( $\kappa_{870} = 0.05 \text{ m}^2 \text{ kg}^{-1}$  for the dust model used in this work),  $\nu_{870} = 343$  GHz, and  $C_\nu(z)$  is a known constant for a given  $\nu$ ,  $\kappa_{870}$ ,  $\beta$ , and  $z$  and  $d_L$ .

The long-wavelength ( $\lambda \gtrsim 200$   $\mu\text{m}$ ) RJ tail of the dust emission, where dust optical depth becomes low, can be well fitted by the above equation. However, equation (6) is also frequently adopted to fit the *full* dust SED, including both the Wien and RJ sides, especially by the studies in the pre-*Herschel* era, when not enough data are available to well cover both sides from the SED peak (Magnelli et al. 2012). The single- $T$  parameter in equation (6) is then often referred

to as the ‘dust temperature’ of the galaxy. However, an effective temperature derived this way should be primarily understood as a fitting parameter and may not correspond to a *physical* temperature (Simpson et al. 2017). In particular, it differs in general from the mass-weighted temperature of dust in a galaxy.

### 3.1.4 Equivalent dust temperature $T_{\text{eqv}}$

We define  $T_{\text{eqv}}$  as the temperature that reproduces the actual IR luminosity for a given broad-band flux (e.g. at 870  $\mu\text{m}$ ) and adopted parametrized functional form of the SED (e.g. OT-MBB). The value of  $T_{\text{eqv}}$  typically depends on both the observing frequency band as well as the SED form (Section 4).

In the specific case of optically thin dust emission, the specific luminosity can be written as

$$\begin{aligned} L_{\nu, \text{OT}}(T, M_{\text{dust}}) &= 4\pi(1+z)^{-1} d_L^2 S_{\nu_0} \\ &= 4\pi \kappa_\nu M_{\text{dust}} B_\nu(T). \end{aligned} \quad (7)$$

By directly integrating the above formula over  $\nu$ , one obtains the total IR luminosity (e.g. Hayward et al. 2011)

$$\begin{aligned} L_{\text{IR, OT}}(T, M_{\text{dust}}) &= \int_0^\infty 4\pi M_{\text{dust}} \kappa_\nu B_\nu(T) d\nu \\ &= 4\pi M_{\text{dust}} \kappa_{\nu_1} \nu_1^{-\beta} \left( \frac{k_B T}{h} \right)^{4+\beta} \left( \frac{2h}{c^2} \right) \\ &\quad \Gamma(4+\beta) \zeta(4+\beta) \\ &= \mathcal{D} M_{\text{dust}} T^{(4+\beta)}, \end{aligned} \quad (8)$$

where  $\mathcal{D}(\kappa_{\nu_1}, \nu_1, \beta)$  is a constant and  $\Gamma$  and  $\zeta$  are Riemann functions.

Combining equations (8) and 6,  $T_{\text{eqv}}$  can now be defined as the temperature satisfying

$$L_{\text{IR}}/S_{\nu_0} = \frac{\mathcal{D} T_{\text{eqv}}^{4+\beta}}{C_\nu(z) B_\nu(T_{\text{eqv}})}. \quad (9)$$

In the RJ regime, where  $B_\nu(T_{\text{eqv}}) = 2\nu^2 k_B T_{\text{eqv}}/c^2$ ,

$$L_{\text{IR}}/S_{\nu_0} \propto T_{\text{eqv}}^{3+\beta}. \quad (10)$$

$T_{\text{eqv}}$  is therefore the temperature that one would need to adopt in order to obtain the correct IR luminosity and match the broad-band flux density under the assumption that the SED has the shape of an OT-MBB function. Of course, the latter assumption is often a poor one and the actual SED shape can differ substantially from an OT-MBB curve. In this case, the equivalent temperature will be different from the mass-weighted temperature. Furthermore, the dust mass that is derived this way [via equation (6) for a given  $T_{\text{eqv}}$  and  $S_{\nu_0}$ ] will then differ from the actual physical dust mass.

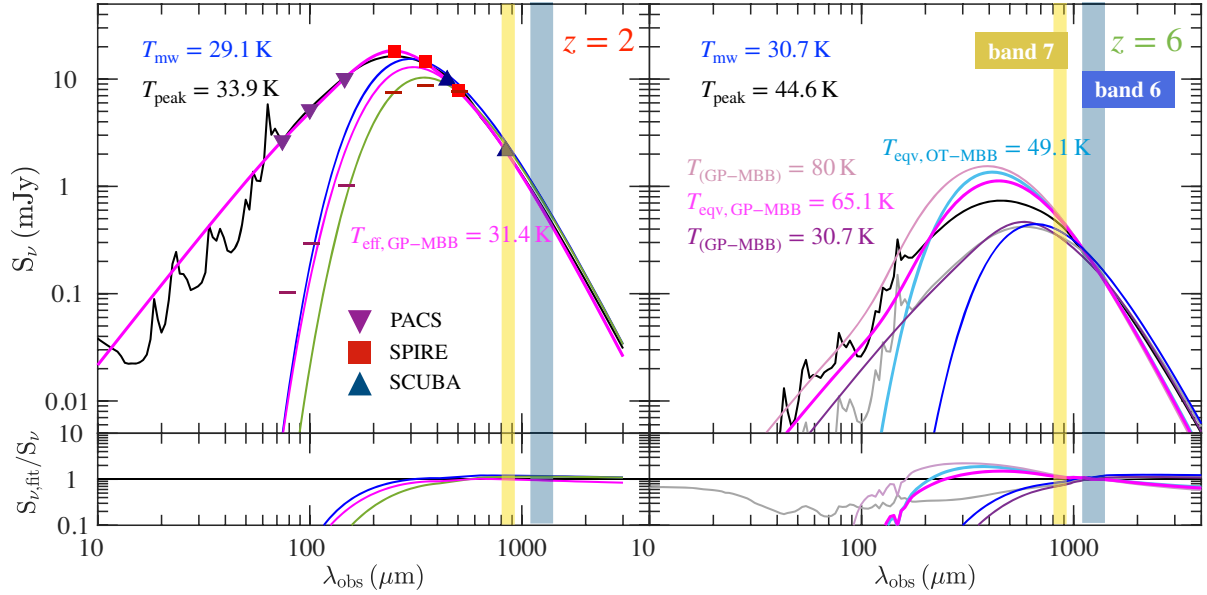
In this paper, we compute  $T_{\text{eqv}}$  based on equation (9) using the actual integrated IR luminosities and 870  $\mu\text{m}$  (1.2 mm) flux densities unless explicitly noted otherwise. For equivalent temperatures based on G-MBB or GP-MBB spectral shapes, we numerically integrate equations (3) and (5) to obtain the IR luminosity for a given dust temperature and dust mass [analogous to equation (8) for the OT-MBB case].

## 3.2 The SEDs of simulated galaxies

In Fig. 3, we show example SEDs of a  $z = 2$  galaxy and a  $z = 6$  galaxy from the MASSIVEFIRE sample. We separately discuss  $z = 2$

<sup>5</sup>Throughout this paper, all  $\nu$  and  $\lambda$  with no subscript stand for rest-frame quantities, while those with ‘o’ are the observed quantities.





**Figure 3.** SKIRT SED (black lines) of selected  $z = 2$  (upper left panel) and  $z = 6$  (upper right panel) MASSIVEFIRE galaxies and SED fitting functions (coloured lines) for the two galaxies. In the upper left panel, the thick magenta line represents the GP-MBB function (equation 5, with  $\alpha = 2.5$ ,  $\beta = 2.0$ , and  $\lambda_1 = 100 \mu\text{m}$ ) that best fits PACS + SPIRE + SCUBA + ALMA photometry calculated from its SKIRT SED. The thin magenta line represents the MBB component of the GP-MBB function. The derived effective temperature  $T_{\text{eff}}$  of the GP-MBB function is 31.4 K. The blue line shows the OT-MBB function (equation 6, with  $\beta = 2.0$ ) with  $T$  being equal to the mass-weighted temperature  $T_{\text{mw}} = 29.1 \text{ K}$  of the galaxy. The green line shows the G-MBB function with the same  $M_{\text{dust}}$  and  $T$  but  $\lambda_1 = 100 \mu\text{m}$ . The optical depth in the G-MBB function results in a lower luminosity-to-mass ratio as well as a longer emission peak wavelength than the OT-MBB function with the same  $M_{\text{dust}}$  and  $T$ . The calculated PACS, SPIRE, and SCUBA flux densities of the galaxy are explicitly marked with the different symbols as labelled, and the horizontal ticks mark the confusion noise limit of the PACS/SPIRE bands. In the upper right panel, we show the GP-MBB (thick salmon, magenta, and purple lines) and OT-MBB (light blue line) functions that are normalized to match the observed flux density at ALMA band 6 (1.2 mm). The magenta and light blue lines correspond to the MBB functions with  $T = T_{\text{eqv}}$  that yield the  $L_{\text{IR}}$  of the  $z = 6$  galaxy. The salmon (purple) line corresponds to a GP-MBB function with  $T > T_{\text{eqv}}$  ( $T < T_{\text{eqv}}$ ), resulting in an over(under)estimate of  $L_{\text{IR}}$ . Like in the upper left panel, we show with blue line the OT-MBB function with  $T = T_{\text{mw}}$  and  $M_{\text{dust}}$  of the selected galaxy. The grey line represents the SED of the  $z = 2$  galaxy that is redshifted to  $z = 6$  and rescaled to match the observed flux density of the  $z = 6$  galaxy at ALMA band 6. In the two upper panels, the golden and grey shaded regions mark ALMA band 7 and 6, respectively. In the lower panels, the coloured lines show the ratio of the flux of the MBB fitting functions [excluding power-law component in equation (5) for the GP cases] to the simulated flux calculated by SKIRT that are shown in the upper panels. An OT-MBB function with  $T_{\text{mw}}$  fits the RJ part of the dust SED quite well, while a GP-MBB function is able to also match the dust SED left of the peak.

and the  $z = 6$  galaxies because the observational strategies for the two epochs are usually different. For  $z = 2$ , an IR-luminous (i.e.  $L_{\text{IR}} \gtrsim 10^{12} L_{\odot}$ ) galaxy may have both *Herschel* coverage at FIR as well as (sub)mm coverage from ground-based facilities (e.g. SCUBA, ALMA, and AzTEC). One can then *derive* the dust temperature ( $T_{\text{peak}}$  or  $T_{\text{eff}}$ ) from the observed FIR-to-mm photometry via SED fitting. At  $z > 4$ , the sources that have a good coverage of the SED peak (via *Herschel* surveys) are currently limited to higher IR luminosity (i.e.  $L_{\text{IR}} \gtrsim 10^{13} L_{\odot}$ ) and the majority are strongly lensed objects (e.g. Weiß et al. 2013; Ivison et al. 2016; Strandet et al. 2016; Miller et al. 2018; Zavala et al. 2018a; Rennehan et al. 2019). Meanwhile, the unprecedented sensitivity of ALMA has allowed us to detect the dust continuum of a growing population of galaxies at these epochs (e.g. Capak et al. 2015; Laporte et al. 2017; Hatsukade et al. 2018; Matthee et al. 2019) that do not have detected *Herschel* counterparts. Most of these observations cover only a single band (typically at ALMA band 6 or 7). Physical properties, such as  $L_{\text{IR}}$  and SFR, are thus often derived based on a single data point at (sub)mm (e.g. Faisst et al. 2017; Schreiber et al. 2018), by *assuming* a dust temperature for the object (e.g. Bouwens et al. 2016; Casey et al. 2018b). This approach is sensible if the adopted dust temperature is close to  $T_{\text{eqv}}$  of the given galaxy (see Section 3.1).

### 3.2.1 Example: the SED of a galaxy at $z = 2$

Fig. 3 shows the SED of a selected  $z = 2$  MASSIVEFIRE galaxy (upper left panel). This galaxy has  $L_{\text{IR}} = 3.0 \times 10^{12} L_{\odot}$ ,  $\text{SFR} = 210 M_{\odot} \text{ yr}^{-1}$ ,  $M_{\text{dust}} = 5.4 \times 10^8 M_{\odot}$ , and  $M_{\star} = 5.3 \times 10^{11} M_{\odot}$ .<sup>6</sup>

We calculate the PACS (70, 100, and 160  $\mu\text{m}$ ) + SPIRE (250, 350, and 500  $\mu\text{m}$ ) + SCUBA-2 (450 and 850  $\mu\text{m}$ ) + ALMA (870  $\mu\text{m}$  and 1.2 mm) broad-band flux densities from the simulated SED. We fit its FIR-to-mm photometry – assuming successful detection at every band, as we show in the left-hand panel that the PACS/SPIRE fluxes of this galaxy are above the confusion noise limit (marked by the horizontal ticks) (Nguyen et al. 2010; Magnelli et al. 2013) and the submm fluxes are above the typical sensitivity limit of SCUBA-2 and ALMA – by a GP-MBB function (with  $\lambda_1 = 100 \mu\text{m}$ ,  $\beta = 2.0$ , and  $\alpha = 2.5$ ) using least- $\chi^2$  method.  $N_{\text{pl}}$  and  $T$  are left as two free parameters for the fitting. The best-fitting GP-MBB function is shown by the thick magenta line. The derived  $T_{\text{eff}}$  is 31.4 K, which

<sup>6</sup>Physical properties of the simulated galaxies reported in this paper are measured using the material within a 30 pkpc kernel around the DM halo centre, i.e. the minimum gravitational potential.



is similar to its mass-weighted temperature ( $T_{\text{mw}} = 29.1$  K).<sup>7</sup> From the best-fitting GP-MBB function (and also the simulated SED),  $T_{\text{peak}}$  is found to be 33.9 K.

For demonstration purpose, we also show with the blue line the exact solution of the OT-MBB function, with  $T = T_{\text{mw}} = 29.1$  K,  $M_{\text{dust}} = 5.4 \times 10^8 M_{\odot}$ ,  $\kappa_{870} = 0.05 \text{ m}^2 \text{ kg}^{-1}$ , and  $\beta = 2.0$ . As expected, the OT-MBB function with a mass-weighted temperature is in very good agreement with the galaxy SED at long wavelength. For this galaxy, at  $\lambda = 100\text{--}650 \mu\text{m}$  [ $\lambda_0 = (300 \mu\text{m})\text{--}(2 \text{ mm})$ ], the difference between the flux of the OT-MBB function and the simulated flux is within 10 per cent (illustrated by the lower left panel). At shorter wavelength, the emission is more tied to the dense, warm dust component in the galaxy, which is poorly accounted for by this OT-MBB function with a mass-weighted temperature. Overall, the OT-MBB function accounts for  $\sim 50$  per cent of  $L_{\text{IR}}$  of the galaxy, and the discrepancy is largely due to the MIR emission.

We also show the effect of optical depth. In the upper left panel, the green line shows the analytic solution from a G-MBB (equation 3) function with the same  $M_{\text{dust}}$  and  $T$  ( $T = T_{\text{mw}} = 29.1$  K), but with a power-law optical depth that equals unity at rest-frame  $\nu_0 = 1.5$  THz, or  $\lambda = 100 \mu\text{m}$ . While the emission looks identical to the optical-thin case (blue line) at long wavelength ( $\lambda_0 > 500 \mu\text{m}$ ), it appears to be lower at shorter wavelength when the effect of optical depth becomes important. The effect of increasing optical depth is that the overall light-to-mass ratio is lower and the emission peak wavelength is longer compared to the optically thin case (cf. Scoville 2013).

### 3.2.2 Example: the SED of a galaxy at $z = 6$

Fig. 3 also shows the SED of a  $z = 6$  MASSIVEFIRE galaxy. This galaxy has lower  $L_{\text{IR}}$  ( $3 \times 10^{11} L_{\odot}$ ) and  $M_{\text{dust}}$  ( $8 \times 10^7 M_{\odot}$ ) compared to the  $z = 2$  galaxy, but interestingly, it has similar  $T_{\text{mw}}$  (30.7 K). The calculated flux densities at ALMA band 7 ( $S_{870 \mu\text{m}}$ ) and 6 ( $S_{1.2\text{mm}}$ ) are 0.44 and 0.23 mJy, respectively. Like the  $z = 2$  galaxy, an OT-MBB function (blue line) with  $M_{\text{dust}}$  and  $T = T_{\text{mw}}$  can well describe the emission of the  $z = 6$  galaxy at long wavelength (for this case,  $\lambda_0 > 1.2 \text{ mm}$ , or rest-frame  $\lambda > 170 \mu\text{m}$ ), but it only accounts for  $\sim 30$  per cent of  $L_{\text{IR}}$ . A larger fraction of the total emission of this  $z = 6$  galaxy originates from the warm dust component.

To estimate  $L_{\text{IR}}$  of a  $z = 6$  galaxy from  $S_{870 \mu\text{m}}$  (or  $S_{1.2\text{mm}}$ ), one often needs an assumed SED function and an assumed  $T_{\text{eqv}}$  for the adopted function. Since it is extremely difficult to constrain the details of SED shape at this high redshift, often a simple OT-MBB or GP-MBB function is used by the observational studies (e.g. Capak et al. 2015; Bouwens et al. 2016; Casey et al. 2018b). As an example, we fit the OT-MBB function to  $S_{1.2\text{mm}}$  of the  $z = 6$  MASSIVEFIRE galaxy with varying  $T$ . We show in the right-hand panel of Fig. 3 the OT-MBB function (with fixed  $\beta = 2.0$ ) that yields the simulated  $L_{\text{IR}}$  with the light blue line. The derived  $T_{\text{eqv}}$  for this function is 49.1 K. This is significantly higher than  $T_{\text{mw}}$ , and as a result, the RJ side of the derived SED of this function appears to be much steeper than the simulated SED. It also poorly fits the simulated SED at wavelength close to  $\lambda_{\text{peak}}$ . The derived  $T_{\text{peak}}$  is therefore very different from the true  $T_{\text{peak}}$  of the simulated SED.

<sup>7</sup>How well  $T_{\text{eff}}$  in the best-fitting GP-MBB function approximates  $T_{\text{mw}}$  depends on its parametrization (see Section 3.1). For instance, increasing  $\lambda_1$  in equation (5) from 100 to 200  $\mu\text{m}$  changes  $T_{\text{eff}}$  from 29.1 to 48.2 K (see also fig. 20 of Casey et al. 2014).

We also fit  $S_{1.2\text{mm}}$  of this galaxy by a GP-MBB function ( $\lambda_1 = 100 \mu\text{m}$ ,  $\beta = 2.0$ ,  $\alpha = 2.5$ ). We show the result for  $T = 30.7$  K (purple line),  $T = 65.1$  (magenta line), and  $T = 80$  K (salmon line). For  $T = T_{\text{mw}} = 30.7$  K, we use the same normalization of the power-law component as for the  $z = 2$  galaxy (upper left panel), so that the SED shape is similar between these two galaxies. For  $T = 65.1$  and 80.0 K, we use the best-fitting normalization factor derived based on the *local* GOALS sample (see table 1 of Casey 2012). We can see that the GP-MBB function appears to better describe the simulated SED shape compared with OT-MBB function, but in order to fit the simulated SED with reasonably good quality, a different choice of  $N_{\text{pl}}$  and  $\lambda_1$  is needed. With  $T = T_{\text{mw}} = 30.7$  K, the GP-MBB function underpredicts the simulated  $L_{\text{IR}}$  ( $3 \times 10^{11} L_{\odot}$ ) by 70 per cent. Using  $T_{\text{eqv, GP-MBB}} = 65.1$  K, this function leads to the right  $L_{\text{IR}}$ . We also show the result for  $T = 80$  K, which overpredicts the  $L_{\text{IR}}$  by about a factor of 2.

In conclusion, we find that an OT-MBB function with a mass-weighted dust temperature well describes the long-wavelength ( $\lambda \gtrsim 200 \mu\text{m}$ ) part of the dust SED, but it does not well account for the Wien side of the SED and leads to significant underestimate of  $L_{\text{IR}}$ . A GP-MBB function can provide high-quality fitting to the simulated SED with good FIR+(sub)mm photometry of galaxy. Using single-band (sub)mm flux density of  $z > 4$  galaxies,  $T_{\text{eqv}}$  is very different from  $T_{\text{mw}}$  of the galaxy. We will discuss  $T_{\text{eqv}}$  for high-redshift galaxies, its evolution with redshift, and its dependence on other galaxy properties in more detail in Section 4.

### 3.3 Comparing simulation to observation

Due to the high confusion noise level of the *Herschel* PACS/SPIRE cameras, most current observational studies on dust temperature at high redshift are limited to the most IR-luminous galaxies in the Universe. For  $z = 2$ , the observations are generally limited to  $L_{\text{IR}} \gtrsim 10^{12} L_{\odot}$ . Applying the powerful stacking technique to the *Herschel* images, it is also possible to probe the fainter regime of a few  $10^{11} L_{\odot}$  at  $z \sim 2$  (e.g. Thomson et al. 2017; Schreiber et al. 2018). Yet another problem with the observational studies is the strong selection bias with flux-limited surveys, meaning that the selected galaxy sample is limited to increasing IR luminosity with redshift. It is therefore non-trivial to disentangle the dependence of dust temperature on redshift and that on other galaxy properties. Using simulated sample, we do not expect to have such a problem.

We start here by comparing the result of the MASSIVEFIRE sample at  $z = 2$  with the observational data from similar redshift. This is where the luminosity ranges of our simulated galaxies share the largest overlap with the current observational data. The selection methods of the quoted data are summarized in Table 1. At higher redshift, the observations are biased to higher  $L_{\text{IR}}$ . In the following section, we will explicitly discuss the redshift evolution of dust temperatures with the MASSIVEFIRE sample.

We present the result in Fig. 4. In the upper panel, we compare the simulations with the observational data of which the (originally *effective*) dust temperature is derived using SED fitting technique and with MBB functions (i.e. equations 3–6), while in the lower panels, we show examples where the dust temperature of both the simulated and observation data is derived using the SED template libraries. In order to make fair comparison among different observations and with the simulation data, we convert all different  $T_{\text{eff}}$  presented in the literature to  $T_{\text{peak}}$  in the upper panel.  $T_{\text{peak}}$  of the simulated galaxies are derived from the best-fitting GP-MBB function (equation 5, with  $\lambda_1 = 100 \mu\text{m}$ ,  $\beta = 2.0$ , and  $\alpha = 2.5$ ) to the FIR-to-mm photometry.

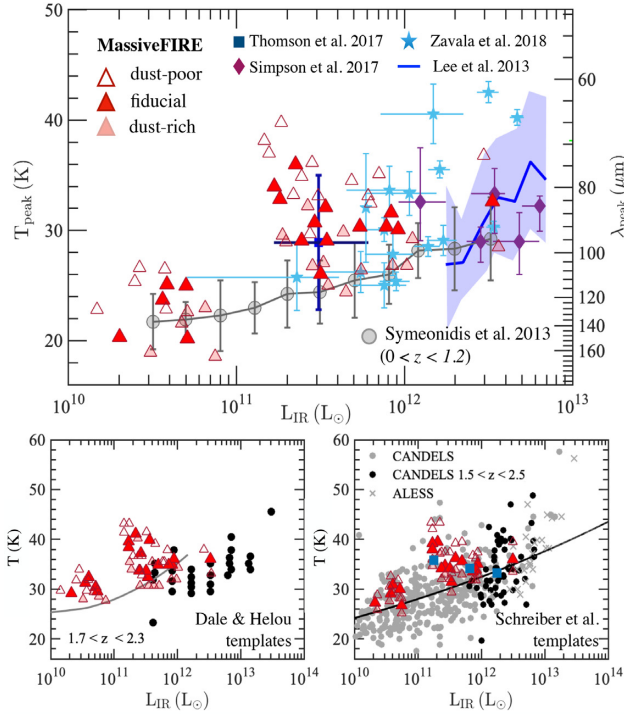
**Table 1.** The selection methods for the observational data presented in Fig. 4.

Data source	Section method
Lee et al. (2013)	<i>Herschel</i> -selected galaxy sample in the COSMOS field with $\geq 5\sigma$ detections in at least two of the five PACS+SPIRE bands and with photometric redshifts (hereafter photo- $z$ ) between 1.5 and 2.0. The $1\sigma$ sensitivity limits are 1.5, 3.3, 2.2, 2.9, and 3.2 mJy in the 100, 160, 250, 350, and 500 $\mu\text{m}$ bands, respectively. Photo- $z$ s are calculated using fluxes in 30 bands that cover the far-UV at 1550 Å to the mid-IR at 8.0 $\mu\text{m}$ .
Symeonidis et al. (2013)	A sample of IR-selected ( <i>Spitzer</i> MIPS 24 $\mu\text{m}$ + <i>Herschel</i> PACS/SPIRE) galaxies at $z = 0-2$ in the COSMOS and GOODS N+S fields. The sample is confined to those 24 $\mu\text{m}$ -detected ( $f_{24} > 30 \mu\text{Jy}$ for GOODS N+S and $f_{24} > 60 \mu\text{Jy}$ for COSMOS) sources that have at least two reliable photometric data points in the two <i>Herschel</i> bands ( $> 3\sigma$ ). 1/3 of the sample have spectroscopic redshifts and the rest photometric redshifts.
Simpson et al. (2017)	SCUBA-2-detected ( $\sigma_{850} = 2.0 \text{ mJy}$ , at $\geq 4\sigma$ ) galaxies in the UKIDSS Ultra Deep Survey (UDS) field with photo- $z$ between 1.5 and 2.5. Photo- $z$ s are determined using 11 bands, covering from the <i>U</i> band to near-IR at 4.5 $\mu\text{m}$ .
Thomson et al. (2017)	535 galaxies detected in the HiZELS at the <i>K</i> band, corresponding to the redshifted wavelength of $\text{H}_{\alpha}$ line at $z = 2.23$ . The sample is confined to those with dust-corrected luminosities $L_{\text{H}\alpha} \geq 2.96 \times 10^{42} \text{ erg s}^{-1}$ , corresponding to an SFR of $\approx 4 M_{\odot} \text{ yr}^{-1}$ .
Zavala et al. (2018b)	SCUBA-2-selected galaxies in the EGS field detected at $> 3.75\sigma$ at 450 and/or 850 $\mu\text{m}$ ( $\sigma_{450} = 1.9$ and $\sigma_{850} = 0.46 \text{ mJy beam}^{-1}$ ). The PACS/SPIRE photometry is obtained from the PACS Evolutionary Probe (PEP; Lutz et al. 2011) and the <i>Herschel</i> Multitiered Extragalactic Survey (HerMES; Oliver et al. 2012) programmes. This sample consists of objects with optical spectroscopic redshift, optical photometric redshift, and FIR photometric redshift estimates.
Magnelli et al. (2014)	The sample consists of the NIR-selected galaxies in the GOODS-N ( $K_s < 24.3$ , down to a $3\sigma$ significance), GOODS-S ( $K_s < 24.3$ , down to a $5\sigma$ significance), and COSMOS ( $K_i < 25$ , down to a $3\sigma$ significance) fields. For each SFR- $M_*$ bin, the mean dust temperature of the galaxies in the bin is derived using their mean PACS+SPIRE flux densities. 29, 26, and 3 per cent of these galaxies have spectroscopic redshift estimates, and the rest have photometric redshift estimates based on the available optical-to-NIR data.
Schreiber et al. (2018)	The sample consists of the $z = 0.3-4$ NIR-selected (down to a $5\sigma$ significance) galaxies in the GOODS-N ( $K_s < 24.5$ ), GOODS-S ( $H < 27.4-29.7$ ), UDS ( $H < 27.1-27.6$ ), and COSMOS ( $H < 27.4-27.8$ and $K_s < 23.4$ for the CANDELS and UVISTA-detected sources, respectively) fields. For $T$ measurement, the galaxies are required to have at least one detection at $\geq 5\sigma$ significance at the <i>Herschel</i> bands on both sides of the peak of the FIR SED. It is also complemented with the $z = 0$ volume-limited sample from the HRS as well as the $z = 2-4$ galaxies in the Extended Chandra Deep Field South (ECDFS) field as part of the ALESS programme. The ALESS sample is selected at 870 $\mu\text{m}$ in the single-dish LABOCA image. For the CANDELS and ECDFS-field galaxies, photo- $z$ s are calculated using the available UV-to-NIR multiwavelength data.

In the upper panel, we show with the blue shaded block the data from the H-ATLAS survey (Lee et al. 2013), which encompasses the high-redshift ( $1.5 < z < 2.0$ ) *Herschel*-selected galaxies in the COSMOS field. The height of the block represents  $1\sigma$  dispersion. We also explicitly show the  $z = 1.5-2.5$  objects from Simpson et al. (2017) (purple diamonds) and Zavala et al. (2018b) (cyan asterisks), which are selected at 850  $\mu\text{m}$  from the deep SCUBA-2 Cosmology Legacy Survey (S2CLS; Geach et al. 2017) probing the Ultra Deep Survey (UDS) and the Extended Groth Strip (EGS) field, respectively. And finally, we present the stacked result by Thomson et al. (2017) (blue square), which is based on a high-redshift ( $\langle z \rangle = 2.23$ ) sample extracted from the High-redshift Emission Line Survey (HiZELS) (Sobral et al. 2013), comprising 388 and 146  $\text{H}_{\alpha}$ -selected star-forming galaxies in the COSMOS and UDS fields, respectively. And for purpose of reference, we show the binned data from Symeonidis et al. (2013) by grey filled circles and error bars, which encompasses an IR-selected sample at  $0.1 < z < 2$  selected from the COSMOS, GOODS-N, and GOODS-S fields. We convert the *effective* dust temperature  $T_{\text{eff}}$  presented in Symeonidis et al. (2013), Simpson et al. (2017), Thomson et al. (2017), and Zavala et al. (2018b) to  $T_{\text{peak}}$ . The relation between  $T_{\text{peak}}$  and  $T_{\text{eff}}$  for the fitting functions that are used by the four studies is plotted in Fig. 5. Simpson et al. (2017) (Thomson et al. 2017) adopt an OT-MBB function (equation 6) with fixed  $\beta = 1.8$  ( $\beta = 1.5$ ), while Symeonidis et al. (2013) (Zavala et al. 2018b) use a G-MBB function (equation 3) with fixed  $\beta = 1.5$  ( $\beta = 1.6$ ) and  $\lambda_1 = 100 \mu\text{m}$ . From Fig. 5, we can see that  $T_{\text{eff}}$  presented in the four studies is higher than  $T_{\text{peak}}$ .

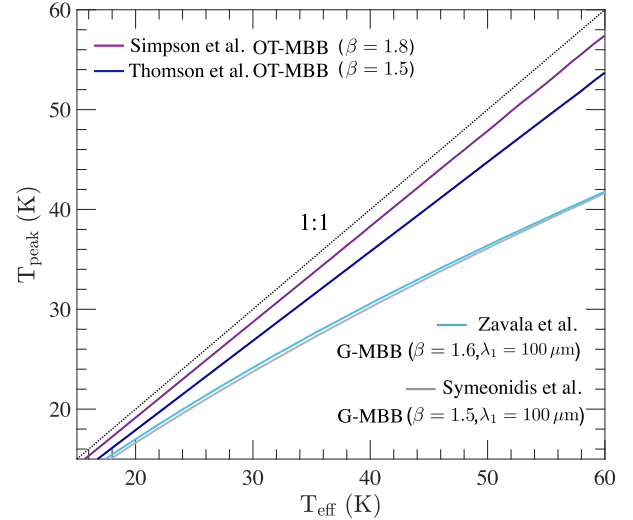
In the lower panels, we compare the simulated result with the observational data from Magnelli et al. (2014) (left) and Schreiber et al. (2018) (right), both of which fit the galaxy photometry to the empirical SED template libraries. In particular, Magnelli et al. (2014) adopt the Dale & Helou (2002) SED template library and determine the temperature for each template by fitting their PACS+SPIRE flux densities with an OT-MBB function with fixed  $\beta = 1.5$  and then finding the  $T_{\text{eff}}$  for the best-fitting OT-MBB function. Their sample comprises near-infrared (NIR)-selected galaxies in GOODS-N, GOODS-S, and COSMOS fields with reliable SFR,  $M_{\text{star}}$ , and redshift estimates. The galaxies are binned in the SFR- $M_{\text{star}} - z$  plane and dust temperatures are inferred using the stacked FIR (100–500  $\mu\text{m}$ ) flux densities of the SFR- $M_{\text{star}} - z$  bins with least- $\chi^2$  method. We show the stacked result for their  $1.7 < z < 2.3$  redshift bin with the black filled dots in the *lower left panel*. For purpose of reference, we also show with the solid grey line the result of a lower redshift bin ( $0.2 < z < 0.5$ ) in the same panel.

In the *lower right panel*, we also compare the simulation to the observational data of Schreiber et al. (2018), of which the galaxy catalogue is based on the CANDELS survey (Grogin et al. 2011; Koekemoer et al. 2011), a  $z = 2-4$  galaxy sample from the ALESS programme (Hodge et al. 2013; Swinbank et al. 2014), as well as the local *Herschel* Reference Survey (HRS; Boselli et al. 2010). The temperature is derived by fitting the PACS+SPIRE photometry to the Schreiber et al. (2018) SED template library, which is constructed based on the Galliano et al. (2011, hereafter G11) library of elementary templates with an assumed power-law



**Figure 4.** Dust temperature versus  $L_{\text{IR}}$  relation of the  $z \sim 2$  galaxies. The red triangles represent the simulated data of the MASSIVEFIRE sample at  $z = 2$ . The unfilled, filled, semitransparent symbols show the result for the dust-poor ( $\delta_{\text{dfr}} = 0.2$ ), fiducial ( $\delta_{\text{dfr}} = 0.4$ ), and dust-rich ( $\delta_{\text{dfr}} = 0.8$ ) models, respectively. In the upper panel, we compare the simulated data to the observational results where dust temperature is derived using the SED fitting technique and with MBB-like functions (equations 3–6). The observation data by Simpson et al. (2017, hereafter S17), Zavala et al. (2018b, hereafter Z18), and the stacked result by Thomson et al. (2017, hereafter T17) are represented by purple diamonds, cyan asterisks, and blue square, respectively. The blue shaded area shows  $1\sigma$  distribution of the compilation of high-redshift COSMOS galaxies by Lee et al. (2013). The grey circles and error bars show the binned result and its  $1\sigma$  distribution of the *Herschel*-selected sample at lower redshift ( $z = 0-1.2$ ) from Symeonidis et al. (2013, hereafter S13). To make fair comparison, we convert  $T_{\text{eff}}$  presented in S13, S17, T17, and Z18 to  $T_{\text{peak}}$ . The relation between  $T_{\text{peak}}$  and  $T_{\text{eff}}$  for each study is shown in Fig. 5. In the lower panels, we show the observational data derived using empirical SED templates. The stacked result by Magnelli et al. (2014) and the compilation by Schreiber et al. (2018) are shown in the left- and right-hand panels, respectively. The solid grey line in the lower left panel represents a second-order polynomial fit to the data points of a lower redshift bin ( $0.2 < z < 0.5$ ). The solid black line in the lower right panel represents the derived  $T-L_{\text{IR}}$  scaling relation by Schreiber et al. (2018) using the combined HRS (Boselli et al. 2010) + CANDELS (Grogin et al. 2011; Hodge et al. 2013) + ALESS (Hodge et al. 2013; Swinbank et al. 2014) samples from local to  $z \sim 4$ . The blue squares show the stacked results for the three luminosity bins at  $z \sim 2$ . The dust temperature in the lower panels is defined using the same method as in Magnelli et al. (2014) and Schreiber et al. (2018). The dust temperature of the  $z = 2$  MASSIVEFIRE sample is in good agreement with the observational data.

distribution of  $U$ . The G11 templates are a set of MIR-to-mm spectra emitted by a uniform dust cloud of  $1 M_{\odot}$  when it is exposed to the Mathis, Mezger & Panagia (1983) interstellar radiation field of a range of  $U$ . The temperature assigned to each Schreiber et al. (2018) template of galaxy SED is the mass-weighted value of the G11 templates being used. We show in the lower right panel the result for the CANDELS sample with the black and grey filled circles.



**Figure 5.**  $T_{\text{peak}}$  versus  $T_{\text{eff}}$  relation of different MBB functions (see also fig. 20 of Casey et al. 2014). The purple, blue, cyan, and grey lines correspond to OT-MBB (equation 6) with  $\beta = 1.8$  (Simpson et al. 2017), OT-MBB with  $\beta = 1.5$  (Thomson et al. 2017), G-MBB (equation 3) with  $\beta = 1.6$  (Zavala et al. 2018b), and G-MBB with  $\beta = 1.5$  (Symeonidis et al. 2013). For the two G-MBB functions,  $\lambda_1 = 100 \mu\text{m}$ . We convert  $T_{\text{eff}}$  reported in the above papers to  $T_{\text{peak}}$  using the  $T_{\text{peak}}$  versus  $T_{\text{eff}}$  relation and show the result in Fig. 4. The black dashed line shows the one-to-one relationship.

The black circles explicitly represent the objects at  $z = 1.5-2.5$ . We also show with blue squares the result of the stacked SEDs for  $z = 1.5-2.5$  derived based on the PACS/SPIRE photometry in the CANDELS sample. The result of the ALESS sample at higher redshift ( $z = 2-4$ ) is shown with grey crosses. The black curve shows the scaling relation  $(T/\text{K}) = 5.57 (L_{\text{IR}}/L_{\odot})^{0.0638}$  that is derived by Schreiber et al. (2018) using the combination of the CANDELS, ALESS, and HRS samples.

For the simulated  $z = 2$  galaxies, we fit their PACS/SPIRE photometry to the Dale & Helou (2002) (as Magnelli et al. 2014) and Schreiber et al. (2018) SED templates using least- $\chi^2$  method and find the temperature associated with the best-fitting template SED as defined in the literature. In other words, the temperature of the MASSIVEFIRE galaxies is not the same in each of the three panels. The temperature derived following the Magnelli et al. (2014) and Schreiber et al. (2018) methods is on average 5.2 and 4.2 K higher than  $T_{\text{peak}}$ , respectively. Comparing the simulated with the observational data, we find an encouragingly good agreement over the common range of  $L_{\text{IR}}$ , with either the observational data derived using SED fitting technique (upper panel), or using SED templates (lower panels). And apart from that,  $T_{\text{peak}}$  of the simulated  $z = 2$  galaxies appear to show no clear correlation with  $L_{\text{IR}}$  in all the three panels, at least at  $L_{\text{IR}} \gtrsim 10^{11} L_{\odot}$ . This is consistent with the recent finding by Schreiber et al. (2018) that the mean dust temperature derived from the stacked SEDs of the three  $L_{\text{IR}}$  bins of their  $z \sim 2$  sample shows almost no correlation over the range of  $1.5 \times 10^{11} - 1.5 \times 10^{12} L_{\odot}$  (blue squares) and lies systematically above the mean temperature of galaxies at lower redshift (black line). This suggests that high-redshift galaxies do not necessarily follow a single, fundamental  $L_{\text{IR}}-T$  scaling relation, which is typically derived using flux-limited observational data across a range of redshift but without much overlap of  $L_{\text{IR}}$  among different redshift bins. We will also show in Section 3.4.3 that the dust temperatures of our MASSIVEFIRE sample increase with redshift



at fixed  $L_{\text{IR}}$  from  $z = 2$  to 6. Ma et al. (2019) also report that the same redshift evolution extends to higher redshift (up to  $z = 10$ ) using a different suite of FIRE simulations.

The observational data show non-trivial scatter, which is particularly clear in the upper and lower right panels. At  $L_{\text{IR}} \approx 3 \times 10^{12} L_{\odot}$ , for instance,  $T_{\text{peak}}$  (upper panel) is observed to be as low as  $\sim 25$  K and as high as  $\sim 45$  K. One possible reason is the intrinsic scatter of  $\delta_{\text{dtr}}$ . We show in Fig. 4 the result for the dust-poor ( $\delta_{\text{dtr}} = 0.2$ ) and dust-rich ( $\delta_{\text{dtr}} = 0.8$ ) models in each panel. The former (latter) show  $\sim 3$  K increase (decrease) of dust temperature(s) compared with the fiducial model ( $\delta_{\text{dtr}}$ ). This difference, however, still appears to be relatively smaller compared to the scatter of the observational data. A larger variance of  $\delta_{\text{dtr}}$  may lead to a larger scatter of temperature. Apart from that, it could also be owing to the variance of the conditions of the ISM structure on the unresolved scale (e.g. compactness and obscurity of the birth-clouds embedding the young stars). We will discuss more about the impact of subgrid models later in Section 5. And finally, given that the *Herschel* cameras have fairly high confusion noise level, and it is rare that one galaxy has full reliable detection at every PACS/SPIRE+SCUBA band, we suggest that both factors can cause non-trivial uncertainty of observational result. Future infrared space telescope (e.g. SPICA; Spinoglio et al. 2017; Egami et al. 2018) spanning similar wavelength range and with higher sensitivity may help improve the constraint near emission peak and hence the observationally derived dust temperatures.

We also note that  $z = 2$  MASSIVEFIRE galaxies appear to show higher dust temperature compared to the lower redshift counterparts in the observed sample, with either the temperature derived using SED fitting (*upper panel*) technique or SED templates (*lower panels*). Observationally, how dust temperature evolves at fixed  $L_{\text{IR}}$  (or  $M_{\text{star}}$ ) from  $z = 0$  to 2 is still being debated (e.g. Hwang et al. 2010; Magdis et al. 2012; Magnelli et al. 2013; Lutz 2014; Magnelli et al. 2014; Béthermin et al. 2015; Kirkpatrick et al. 2017; Schreiber et al. 2018). Uncertainties can potentially arise from selection effects (surveys at certain wavelengths preferentially select galaxies of warmer/colder dust) (e.g. Magdis et al. 2010; Hayward et al. 2011; McAlpine et al. 2019) and inconsistency in derivation of dust temperature. The dust temperature of galaxies in this redshift regime ( $z < 2$ ) is beyond the scope of this paper.

### 3.4 The role of dust temperature in scaling relationships

The scaling relationships of dust temperature against other dust/galaxy properties (such as total IR emission, sSFR, and etc.) have been extensively studied in the past decade because of the significant boost of the number of detected high-redshift dusty star-forming galaxies by *Herschel*, SCUBA, and ALMA. We now have statistically large sample for revealing and studying the various scaling relationships of dust temperature. Here in this section, we show the result of the MASSIVEFIRE sample at  $z = 2-6$ , discuss the physical interpretation of the scaling relations, and specifically examine how each scaling relation differs by using different dust temperatures ( $T_{\text{mw}}$  versus  $T_{\text{peak}}$ ).

#### 3.4.1 $S \propto MT$ (optically thin regime)

As mentioned above, the long-wavelength RJ tail can be well described by a single- $T$  OT-MBB function. This is a direct consequence of the rapid power-law decline of the dust opacity with wavelength as well as the fact that the coldest dust dominates

the mass budget (e.g. Dunne & Eales 2001; Harvey et al. 2013; Lombardi et al. 2014; Utomo et al. 2019). At very long wavelength, the flux is only linearly dependent on  $T$  in the RJ tail, and therefore the overall shape of the SED on the RJ side is largely set by the temperature of the mass-dominating cold dust. Hence, it has been proposed that the flux density originating from the optically thin part of the RJ tail can be used as an efficient measure for estimating dust and gas mass (by assuming a dust-to-gas ratio) of massive high-redshift galaxies (e.g. Magdis et al. 2012; Scoville et al. 2014; Groves et al. 2015; Scoville et al. 2016; Hughes et al. 2017; Liang et al. 2018; Privon, Narayanan & Davé 2018; Kaasinen et al. 2019). Given the high uncertainties of the traditional CO methods and their long observing time, this approach represents an important alternative strategy for gas estimate (e.g. Schinnerer et al. 2016; Scoville et al. 2017b; Harrington et al. 2018; Janowiecki et al. 2018; Cochrane et al. 2019; Wiklind et al. 2019).

The RJ approach benefits from the effect of ‘negative  $K$ -correction’. Equation (6) can be rewritten as (e.g. Scoville et al. 2016)

$$\begin{aligned} S_{\nu_0}(T) &= \frac{(1+z)}{d_L^2} 2k_B \kappa_{\nu}(\nu/c)^2 \Gamma_{\text{RJ}}(\nu_0, T, z) M_{\text{dust}} T \\ &= \psi(z) \Gamma_{\text{RJ}} M_{\text{dust}} T, \end{aligned} \quad (11)$$

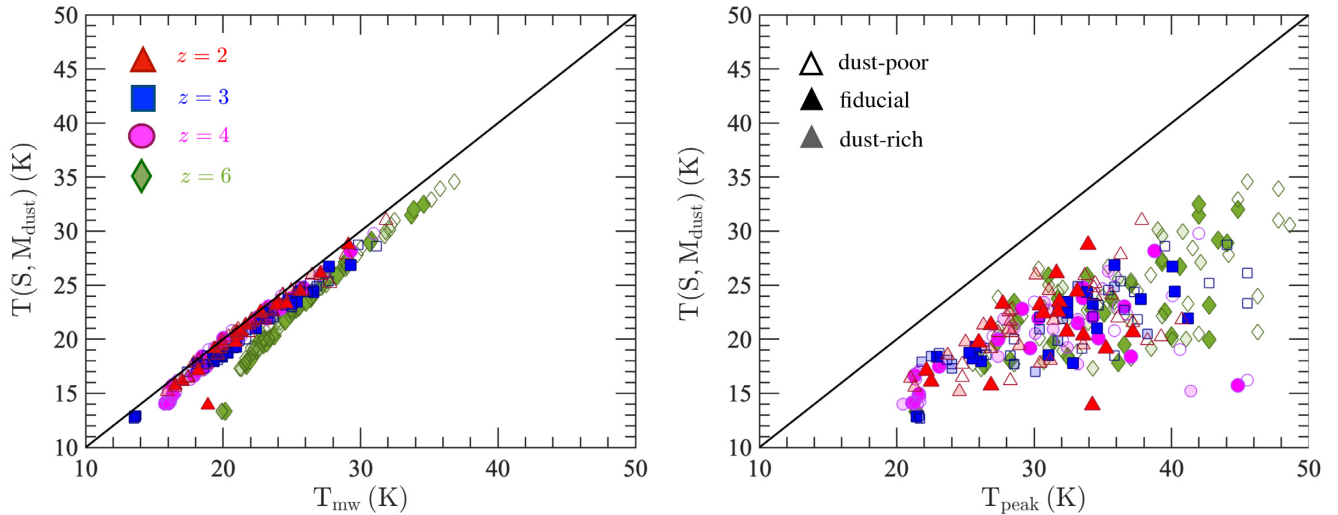
where  $\Gamma_{\text{RJ}}$  is the RJ correction function that accounts for the departure of the Planck function from RJ approximate solution in the rest frame, and  $\psi(z)$  has the unit of  $\text{mJy } M_{\odot}^{-1} \text{ K}^{-1}$ . For given  $\nu_0$ ,  $\kappa_{\nu}(\nu/c)^2$  scales as  $(1+z)^4$  ( $\beta = 2.0$ ). On the other hand,  $(1+z)d_L^{-2}$  and  $\Gamma_{\text{RJ}}$  decline with redshift. The former term roughly scales as  $(1+z)^{-2}$ , while how  $\Gamma_{\text{RJ}}$  evolves with redshift depends on both  $\nu_0$  and  $T$ . The rise of  $\kappa_{\nu}(\nu/c)^2$  with redshift can roughly cancel out or even reverse the decline of the other two components at  $z \gtrsim 1$ , with typical  $T$  of galaxies and (sub)mm bands. For example, with  $T = 25$  K and ALMA band 6,  $\psi \Gamma_{\text{RJ}}$  stays about a constant from  $z = 2-6$ , while with ALMA band 7,  $\psi \Gamma_{\text{RJ}}$  declines *only* by less than a factor of 2 over the same redshift range (see fig. 2 of Scoville et al. 2016). (Sub)mm observations are therefore powerful for unveiling high-redshift dusty star-forming galaxies. In the RJ regime ( $h\nu \ll k_B T$ ),  $\Gamma_{\text{RJ}} \approx 1$  and  $S$  scales linearly to  $M_{\text{dust}} T$  at a given redshift.

The RJ approach relies on an assumed dust temperature. The proper temperature,  $T$ , needed for inferring dust (and gas) masses can be obtained from solving equation (11), given  $S_{\nu_0}$ ,  $M_{\text{dust}}$ , and  $z$ . This required  $T$  value is close to the mass-weighted dust temperature, for galaxies from  $z = 2$  to 6, and with varying  $\delta_{\text{dtr}}$ , see Fig. 6. The difference between these two temperatures is typically less than 0.03 dex. This again confirms that a single- $T$  OT-MBB function well describes the emission from the optically thin RJ tail.

However, using  $T_{\text{peak}}$  will lead to a poor constraint on  $M_{\text{dust}}$  and therefore gas mass of galaxy. First of all, it is systematically higher than  $T_{\text{mw}}$ , and therefore can cause systematically underestimate of  $M_{\text{dust}}$ . Secondly, there seems to be no strong correlation between  $T_{\text{mw}}$  and  $T_{\text{peak}}$  by comparing the *left-* and *right-hand panels*. So even by using  $T_{\text{peak}}$  to infer  $T_{\text{mw}}$  will produce systematic error. We will discuss the discrepancy between  $T_{\text{peak}}$  and  $T_{\text{mw}}$  in more detail in the later sections. *Using other effective temperatures that have strong correlation with  $T_{\text{peak}}$  will be problematic as well.*

#### 3.4.2 The $L_{\text{IR}}$ versus $MT^{4+\beta}$ relation

The scaling relation  $L_{\text{IR}} \propto M_{\text{dust}} T^{(4+\beta)}$ , which is frequently been adopted by many studies to probe and obtain useful physical insights for the star-forming conditions of the IR-luminous sources owing to



**Figure 6.** Relation of the temperature needed for dust mass estimate (calculated from equation 11) against  $T_{\text{mw}}$  (left-hand panel) and  $T_{\text{peak}}$  (right-hand panel) of the MASSIVEFIRE sample at  $z = 2$  (red triangles),  $z = 3$  (blue squares),  $z = 4$  (magenta circles), and  $z = 6$  (green diamonds). For the  $z = 2$ –4 galaxies, the flux density for mass estimate is measured at ALMA band 7 ( $\lambda_o = 870 \mu\text{m}$ ), while for the  $z = 6$  galaxies (green), it is measured at ALMA band 6 ( $\lambda_o = 1.2 \text{ mm}$ ) so as to ensure the rest-frame wavelength is on the optically thin part of the RJ tail. The unfilled, filled, and semitransparent symbols represent the result for  $\delta_{\text{dzt}} = 0.2, 0.4$ , and  $0.8$ , respectively. The solid diagonal line marks the one-to-one locus.  $T_{\text{mw}}$  is the temperature needed for estimating dust mass using the RJ-tail approach.  $T_{\text{peak}}$  is a poor proxy for this temperature.

its simplicity, is derived under the assumption of the optically thin approximation (equation 8).

The temperature in the above scaling relation is a measure of the luminosity per unit dust mass and often viewed as a proxy for the internal radiative intensity. Yet, it is *not* obvious how this temperature parameter [i.e.  $\sim (L_{\text{IR}}/M_{\text{dust}})^{1/6}$ ] is related to the physical,  $T_{\text{mw}}$  or the observationally accessible  $T_{\text{peak}}$ .

We show in Fig. 7 the scaling relation of the light-to-mass ratio,  $L_{\text{IR}}/M_{\text{dust}}$  against  $T_{\text{mw}}$  (left-hand panel) as well as  $T_{\text{peak}}$  (right-hand panel) for the MASSIVEFIRE sample at  $z = 2$ –6, and we explicitly present the result for the fiducial (filled symbols), dust-poor (unfilled symbols) and dust-rich (semitransparent symbols) cases.

In general, galaxy having higher dust temperature (both  $T_{\text{mw}}$  and  $T_{\text{peak}}$ ) emits more IR luminosity per unit dust mass. Focusing at first on  $T_{\text{mw}}$  (left-hand panel), we see that  $L_{\text{IR}}/M_{\text{dust}}$  of the MASSIVEFIRE galaxies appears to be systematically higher than from a simple single- $T$  OT-MBB function (equation 7), which is indicated by the solid black line in both the panels. The offset ( $\sim 0.3$  dex) between the simulated result and the analytic solution is due to the higher emissivity of the dense, warm dust in vicinity of the star-forming regions (see the lower panels of Fig. 2), which accounts for a small fraction of the total dust mass but has strong emission, and shapes the Wien side of the overall SED of galaxy.

With all the galaxies from  $z = 2$  to 6, we find that  $L_{\text{IR}}/M_{\text{dust}}$  scales to  $\approx T_{\text{mw}}^{5.4}$ . This is slightly flatter than the analytic solution derived using a single-temperature, optically thin MBB function, i.e.  $L_{\text{IR,OT}}/M_{\text{dust}} \propto T^6$  (equation 8, with  $\beta = 2.0$ ). We understand the shallower slope as an optical depth effect. In the optically thin regime ( $\tau \ll 1$ ),  $L/M \propto (1 - e^{-\tau})/\tau \approx 1$ , while in the optically thick regime ( $\tau \gg 1$ ),  $L/M \propto \tau^{-1}$  (equation 4). In the optically thick regime,  $L_{\text{IR}}/M_{\text{dust}}$  therefore decreases with increasing  $\tau$ . Galaxies of higher  $T_{\text{mw}}$  are more dust-rich (Section 3.4.3) and their star-forming regions tend to be more optically thick, resulting in a flattening of the scaling relation.

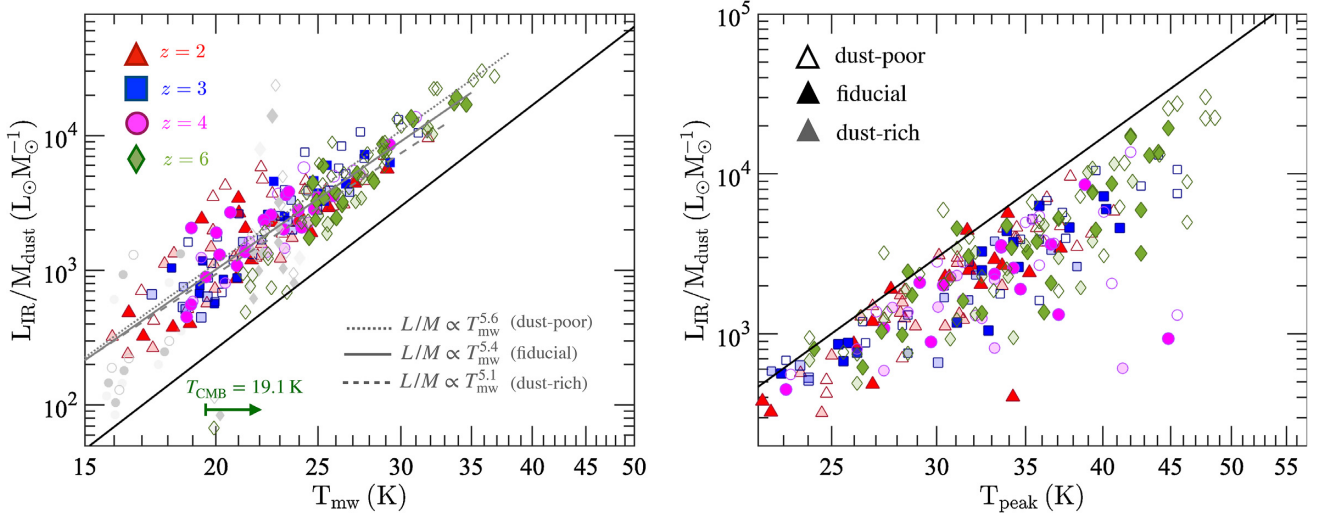
Comparing the dust-poor (dust-rich) models with the fiducial case, the median of  $T_{\text{mw}}$  is higher (lower) by  $0.84$  ( $1.70$ ) K. This is

due to the optical depth effect. By reducing the amount of dust, the chance of receiving a short-wavelength photon increases because the optical depth from the emitting sources decreases. Therefore, dust is expected to be heated to higher temperature to balance the increased amount of absorption. Apart from that,  $\delta_{\text{dzt}}$  also mildly affects the normalization of the  $L_{\text{IR}}/M_{\text{dust}}$  versus  $T_{\text{mw}}$  relation. The dust-poor (dust-rich) case shows about  $0.13$  ( $0.06$ ) dex higher (lower)  $L_{\text{IR}}/M_{\text{dust}}$ , on the average, than the fiducial case, indicating a high (lower) luminosity emitted per unit dust mass. This is because a larger (reduced) mass fraction of the total dust is heated by (can actually ‘see’) the hard UV photons emitted from the young stars due to the reduced optical depth (Scoville 2013; Scoville et al. 2016). This dust component can be efficiently heated to a temperature much higher than the mass-weighted average of the bulk (Harvey et al. 2013; Lombardi et al. 2014; Broekhoven-Fiene et al. 2018), and has a much higher  $L/M$  ratio than the rest.

$T_{\text{peak}}$  (right-hand panel) also shows a positive correlation with  $L_{\text{IR}}/M_{\text{dust}}$ , although the strength of correlation is relatively weaker than that of  $T_{\text{mw}}$  ( $\rho = 0.81$  versus  $0.91$ , where  $\rho$  is the Spearman rank correlation coefficient). Besides,  $T_{\text{peak}}$  also shows larger scatter than  $T_{\text{mw}}$ . The  $1\sigma$  dispersion of  $L_{\text{IR}}/M_{\text{dust}}$  at fixed  $T_{\text{peak}}$  is  $0.21$  dex, which is higher than  $0.14$  dex at fixed  $T_{\text{mw}}$ . This means that  $T_{\text{peak}}$  has relatively lower power to predict the luminosity-to-dust-mass ratio. Furthermore,  $T_{\text{peak}}$  is also more affected by a change of  $\delta_{\text{dzt}}$ . The median  $T_{\text{peak}}$  of the dust-poor (dust-rich) case is  $2.49$  ( $2.63$ ) K higher (lower) than the fiducial model, which is more than the change of  $T_{\text{mw}}$  with  $\delta_{\text{dzt}}$ . This is because  $T_{\text{peak}}$  is more sensitive to the mass fraction of ISM dust that is efficiently heated to high temperature by the hard UV photons emitted from young stars (see also Faisst et al. 2017).

### 3.4.3 $L_{\text{IR}}$ versus $T$ relation

The dust temperature versus total IR luminosity is one most extensively studied scaling relations. We have shown in Section 3.2



**Figure 7.** Relation of  $L_{\text{IR}}/M_{\text{dust}}$  against  $T_{\text{mw}}$  (left-hand panel) and  $T_{\text{peak}}$  (right-hand panel) of the MASSIVEFIRE sample at  $z = 2-6$ . The results for the fiducial, dust-poor, and dust-rich cases are shown with unfilled, filled, and semitransparent symbols, respectively. In the left-hand panel, the dotted, solid, and dashed grey lines represent the best-fitting power-law scaling relation for the dust-rich, fiducial, and dust-poor cases, respectively. Those galaxies of which  $T_{\text{mw}}$  is strongly affected by CMB heating, i.e.  $T_{\text{mw}} - T_{\text{CMB}}(z) < 5$  K, are coloured by grey. They are excluded from the power-law fitting. The dust-rich (poor) case exhibits a flatter (steeper)  $L_{\text{IR}}/M_{\text{dust}}$  versus  $T_{\text{mw}}$  scaling relation compared with the fiducial model. The solid black line in each panel represents the expected analytic scaling using the optically thin MBB function (equation 6), with the dust emissivity spectral index  $\kappa_{870} = 0.05 \text{ m}^2 \text{ kg}^{-1}$ .

that our simulations have successfully produced the result at  $z = 2$  for galaxies that are in good agreement with the recent observational data at similar luminosity range. Here in this section, we focus on the evolution of dust temperature up to higher redshift. One major problem with the current observational studies on the  $T$ - $L$  scaling is the selection effects of the flux-limited FIR samples that have been used to probe such relation. Higher redshift sample is biased towards more luminous systems (Madau & Dickinson 2014). How dust temperature evolves at fixed luminosity is still being routinely debated (see e.g. Magdis et al. 2012; Symeonidis et al. 2013; Magnelli et al. 2014; Béthermin et al. 2015; Ivison et al. 2016; Schreiber et al. 2018; Casey et al. 2018a). We present the result using our sample with  $L_{\text{IR}} \approx 10^9 - 2 \times 10^{12} L_{\odot}$  from  $z = 2$  to 6. For  $z = 3-6$ , there is no current data available that we can make direct comparison to at similar  $L_{\text{IR}}$  of our sample. Future generation of space infrared telescope, such as SPICA, can probe similar regime of IR luminosity at these epochs.

We present the temperature versus luminosity relation of the MASSIVEFIRE galaxies at  $z = 2-6$  in Fig. 8. In the upper and lower left panels, we show  $T_{\text{peak}}$  versus  $L_{\text{IR}}$  and  $T_{\text{mw}}$  versus  $L_{\text{IR}}$  relations, respectively.

Focusing at first on  $T_{\text{peak}}$  versus  $L_{\text{IR}}$  relation (upper left), we find a noticeable increase of  $T_{\text{peak}}$  with redshift at fixed  $L_{\text{IR}}$ , albeit with large scatter at each redshift. Looking at the most luminous galaxy at each redshift, we see that  $T_{\text{peak}}$  increases from about 34 K at  $z = 2$  to  $\sim 43$  K at  $z = 6$  for the fiducial dust model ( $\delta_{\text{dfr}} = 0.4$ ). With all the luminous galaxies with  $L_{\text{IR}} > 10^{11} L_{\odot}$ , we fit the evolution of  $T_{\text{peak}}$  with redshift as a power law and obtained

$$\log \left( \frac{T_{\text{peak}}(z)}{25 \text{ K}} \right) = (-0.02 \pm 0.06) + (0.25 \pm 0.09) \log(1+z). \quad (12)$$

This result is in good quantitative agreement with the recent observational finding by Ivison et al. (2016) and Schreiber et al.

(2018), although they use more IR-luminous sample at similar redshift range.

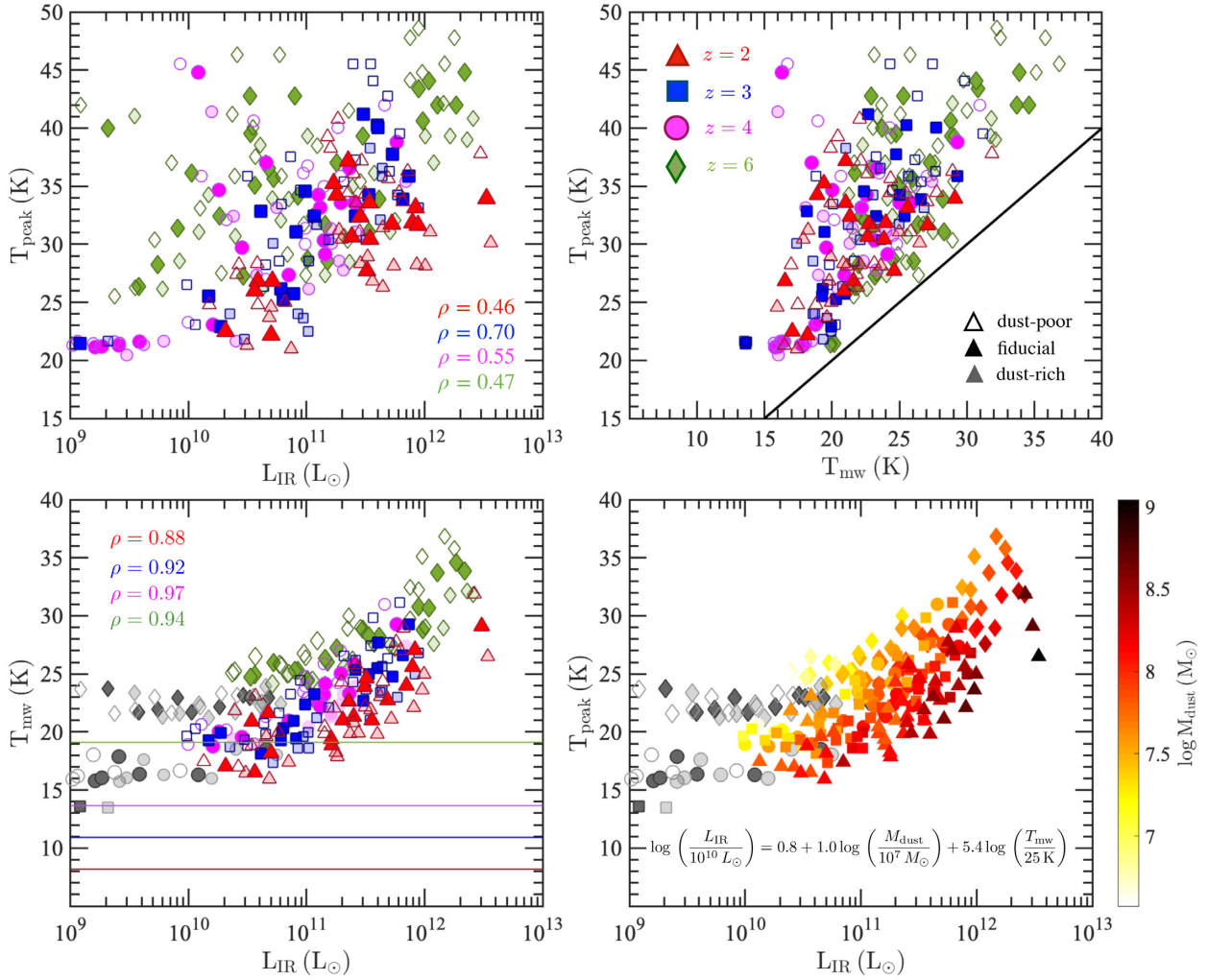
For each redshift, there is also a mild trend of declining  $T_{\text{peak}}$  with decreasing  $L_{\text{IR}}$  over the three orders of magnitude of  $L_{\text{IR}}$  being considered. For instance,  $T_{\text{peak}}$  of the  $z = 6$  galaxies at  $L_{\text{IR}} = 10^{10} L_{\odot}$  is about 32 K, which is about 10 K lower than the value at  $L_{\text{IR}} = 10^{12} L_{\odot}$ , and is similar to the value of the brightest objects at  $z = 3$  and 4. We find some faint objects at  $\sim 10^{10} L_{\odot}$  whose  $T_{\text{peak}}$  is as low as  $\sim 20$  K. We also note that the scatter of  $T_{\text{peak}}$  could be very large at the faint end even with the simple fiducial dust model. At  $z = 4$ , some objects could be as hot as  $\sim 40$  K, while some could be as cold as  $\sim 20$  K. This large scatter is mainly driven by the difference of sSFR among those galaxies, which we will discuss in more detail in the following section.

With such large scatter, the correlation between  $T_{\text{peak}}$  and  $L_{\text{IR}}$  appears to be fairly weak. The Spearman correlation coefficient ( $\rho$ ) of the  $T_{\text{peak}}$  versus  $L_{\text{IR}}$  relation at individual redshift ranges from 0.46 to 0.70 at the redshifts being considered. For the  $z = 2$  sample, there is no noticeable correlation at  $L_{\text{IR}} > 10^{11} L_{\odot}$ .

On the other hand,  $T_{\text{mw}}$  exhibits a tighter correlation with  $L_{\text{IR}}$  (lower left panel) ( $\rho$  ranging from 0.88 to 0.97), with an increase of the normalization of the  $L_{\text{IR}}-T_{\text{mw}}$  relation with redshift. The increase of  $T_{\text{mw}}$  with redshift at fixed  $L_{\text{IR}}$  is clearly less prominent than  $T_{\text{peak}}$ . At  $L_{\text{IR}} \approx 10^{12} L_{\odot}$ , for example,  $T_{\text{mw}}$  increases from  $\sim 27$  K at  $z = 2$  to only  $\sim 32$  K at  $z = 6$ . The CMB heating sets a temperature floor for  $T_{\text{mw}}$  at the low-luminosity end.

The evolution of the  $T_{\text{mw}}$  versus  $L_{\text{IR}}$  scaling is driven by  $M_{\text{dust}}$ . At fixed  $L_{\text{IR}}$ , galaxies at higher redshift have lower  $M_{\text{dust}}$ . This can be clearly seen from the lower right panel, where we colour the same data as in the lower left panel by  $M_{\text{dust}}$  of galaxy. There is clear sign of anticorrelation between  $T_{\text{mw}}$  and  $M_{\text{dust}}$  at fixed  $L_{\text{IR}}$  (see also Hayward et al. 2012; Béthermin et al. 2015; Safarzadeh et al. 2016; Faist et al. 2017; Kirkpatrick et al. 2017). Applying multivariable linear regression analysis to the  $z = 2-6$  galaxies, excluding those





**Figure 8.** *Upper left:*  $T_{\text{peak}}$  versus  $L_{\text{IR}}$  relation of the MASSIVEFIRE galaxies at  $z = 2-6$ . *Upper right:*  $T_{\text{peak}}$  versus  $T_{\text{mw}}$  relation. *Lower panels:*  $T_{\text{mw}}$  versus  $L_{\text{IR}}$  relation. In the *left-hand panel*, galaxies are coloured by their redshift, while in the *right-hand panel*, they are coloured by  $M_{\text{dust}}$ . The galaxies that are strongly affected by the heating of the CMB background [i.e.  $T_{\text{mw}} \lesssim T_{\text{CMB}}(z) + 5$  K] are coloured by grey. The horizontal solid lines in the *lower left panel* represent the CMB temperature at each redshift. In the upper panels and the *lower left panel*, the filled, unfilled, and the semitransparent symbols represent the fiducial, dust-poor, and dust-rich models, respectively. The data from all three dust models are included in the lower right panel.

that are strongly affected by the heating of the CMB background [i.e.  $T_{\text{mw}} \lesssim T_{\text{CMB}}(z) + 5$  K], we obtain the scaling relation

$$\log \left( \frac{L_{\text{IR}}}{10^{10} L_{\odot}} \right) = (0.81 \pm 0.07) + (1.01 \pm 0.06) \log \left( \frac{M_{\text{dust}}}{10^7 M_{\odot}} \right) + (5.40 \pm 0.36) \log \left( \frac{T_{\text{mw}}}{25 \text{ K}} \right),$$

or  $L_{\text{IR}} \propto M_{\text{dust}}^{5.4} T_{\text{mw}}^{2.0}$  ( $\beta = 2.0$ ). (13)

It appears to be shallower than the classical  $L_{\text{IR}} \propto M_{\text{dust}}^{T^{(4+\beta)}}$  ( $\beta = 2.0$  for our adopted dust model, cf. Fig. 1) relation derived based on the optically thin approximation. We will discuss in Section 5 about using this scaling relation to estimate  $M_{\text{dust}}$  and  $T_{\text{mw}}$  when only single data point is available at FIR-to-mm wavelengths.

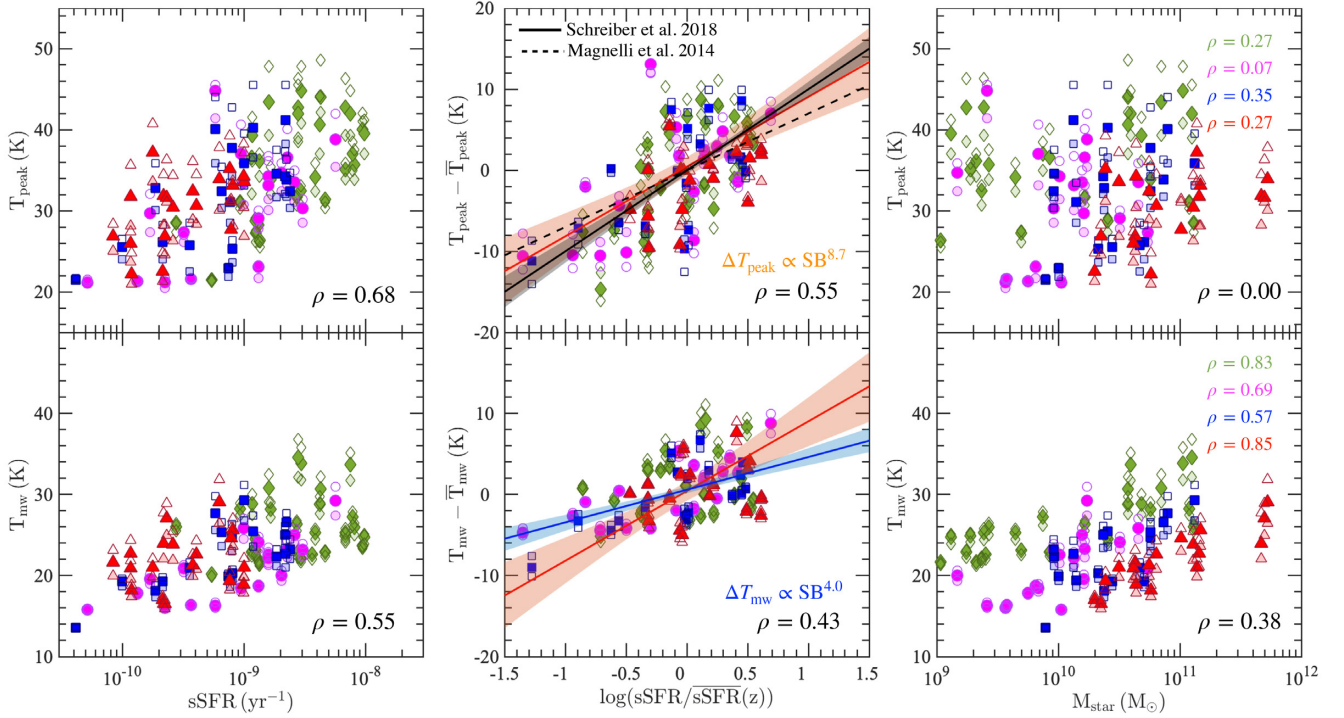
### 3.4.4 sSFR versus $T$ relation

The sSFR versus dust temperature relation is one other frequently studied scaling relation that provides useful physical insights to dust

temperature and is complementary to the  $L_{\text{IR}}$  versus temperature relation.

In Fig. 9, we show the relation of dust temperature against  $\text{sSFR} = \text{SFR}_{20 \text{ Myr}}/M_{\text{star}}$  for the MASSIVEFIRE sample at  $z = 2-6$  in the *left-hand panels*. We present the result for  $T_{\text{peak}}$  and  $T_{\text{mw}}$  in the upper and lower *left panels*, respectively.

The dust temperatures are positively correlated with sSFR ( $\rho = 0.68$  for the sSFR versus  $T_{\text{peak}}$  relation and  $\rho = 0.55$  for the sSFR versus  $T_{\text{mw}}$  relation). Galaxies at higher redshift have, on average, higher sSFR, which is a direct consequence of the evolution of the star formation main sequence. SFR is a proxy for the internal radiative intensity (most UV emission originates from the young stellar populations in the galaxies), and  $M_{\text{dust}}$  is about linearly scaled to  $M_{\text{star}}$  in the MASSIVEFIRE galaxies; the sSFR ( $\sim \text{SFR}/M_{\text{dust}}$ ) can be viewed as a proxy for the total energy input rate per unit dust mass. It is therefore expected that to first order, dust temperature is positively correlated with sSFR of galaxies. This is indeed what we can see from both of the *left-hand panels* of Fig. 9. For instance, the  $z = 2$  galaxies (red) have a median sSFR of  $3 \times 10^{-9} \text{ yr}^{-1}$  and



**Figure 9.** Relation of dust temperature against sSFR (left column), SB =  $\text{sSFR}/\sqrt{\text{sSFR}(z)}$  (middle column), and  $M_{\text{star}}$  (right column) of the MASSIVEFIRE galaxies at  $z = 2$  (red triangles),  $z = 3$  (blue squares),  $z = 4$  (magenta circles), and  $z = 6$  (green diamonds). The results of  $\delta_{\text{dzt}} = 0.4, 0.8$ , and  $0.2$  are shown with filled, semitransparent, and unfilled symbols, respectively. We show the result with  $T_{\text{peak}}$  and  $T_{\text{mw}}$  in the upper and lower panels, respectively. The solid and dashed lines in the upper middle panel represent the observed scaling relation by Schreiber et al. (2018) and Magnelli et al. (2014), respectively. The orange and blue lines in the middle panels show the best-fitting line for the MASSIVEFIRE sample. The former and latter correspond to  $T_{\text{peak}}$  and  $T_{\text{mw}}$ , respectively. The shaded areas represent the 95 per cent (i.e.  $2\sigma$ ) confidence interval of each scaling relation.  $T_{\text{peak}}$  exhibits relatively stronger correlation with sSFR than  $T_{\text{mw}}$ , but weaker correlation with  $M_{\text{star}}$ .

median  $T_{\text{mw}} = 20$  K ( $T_{\text{peak}} = 30$  K). Both sSFR and dust temperature (both  $T_{\text{peak}}$  and  $T_{\text{mw}}$ ), on average, increase with redshift. The  $z = 6$  sample (green) has a median sSFR of  $2 \times 10^{-8} \text{ yr}^{-1}$  and median  $T_{\text{mw}} = 26$  K ( $T_{\text{peak}} = 37$  K).

The correlation persists when focusing on each individual redshift. In the middle panels, we show the result when both temperature and sSFR are normalized by the median value of the whole sample ( $\bar{T}_{\text{mw}}$  or  $\bar{T}_{\text{peak}}$ ,  $\text{sSFR}$ ) at each different redshift. With  $T_{\text{peak}}$  (upper middle panel), the simulated galaxies, including all objects at  $z = 2-6$ , exhibit a positive correlation ( $\rho = 0.55$ ) between starburstiness<sup>8</sup> [i.e.  $\text{SB} = \text{sSFR}/\sqrt{\text{sSFR}(z)}$ ] and normalized  $T_{\text{peak}}$ . The derived scaling relation (solid orange line) is in good qualitative agreement with the recent observations by Magnelli et al. (2014) (dotted black line) and Schreiber et al. (2018) (solid black line), despite that both studies include samples at lower redshifts ( $z < 2$ ) that our simulations do not probe. We also find that compared to  $L_{\text{IR}}$ ,  $T_{\text{peak}}$  is more strongly correlated with sSFR at each given redshift, which is in agreement with the previous finding by Magnelli et al. (2014) (see also Lutz 2014).

However, due to the inhomogeneity of dust distribution in galaxies and the complexity in star-dust geometry, the radiative energy emitted from the young stellar populations is not expected to evenly heat the ISM dust in the galaxy. Most of the UV

photons are absorbed by the dense dust cloud in vicinity of the young star-forming regions, while the majority of the dust in the ISM is heated by the old stellar populations with more extended distribution, as well as the secondary photons re-emitted from the dust cloud near the young star clusters. For such reason,  $T_{\text{peak}}$  is expected to be more sensitive to the emission from the warm dust component, which is more closely tied to the young star clusters, while  $T_{\text{mw}}$  is determined by the cold dust component and therefore can be relatively less sensitive to the sSFR of galaxy than  $T_{\text{peak}}$ .

This indeed can be seen from comparing the *upper* and *lower middle panels* of Fig. 9. First of all,  $\Delta T_{\text{peak}}$  ( $T_{\text{peak}} - \bar{T}_{\text{peak}}$ ) shows a relatively stronger correlation with SB than  $\Delta T_{\text{mw}}$  ( $T_{\text{mw}} - \bar{T}_{\text{mw}}$ ). With all the  $z = 2-6$  MASSIVEFIRE galaxies, the Spearman correlation coefficient of the  $\Delta T_{\text{peak}}$  versus SB scaling is  $\rho = 0.55$ , while that of the  $\Delta T_{\text{mw}}$  versus SB scaling is  $\rho = 0.43$ . Secondly, over about two orders of magnitude of SB ( $\sim 0.1-10$ ), the scaling relation with  $\Delta T_{\text{peak}}$  is relatively steeper,

$$\Delta T_{\text{peak}} \propto \text{SB}^{8.61 \pm 1.38} \text{ versus } \Delta T_{\text{mw}} \propto \text{SB}^{4.03 \pm 0.48}. \quad (14)$$

This is because the UV photons from the young star clusters preferentially heat the dense dust cloud in the neighbourhood to high temperature, which boosts the MIR emission and helps shift the SED peak to shorter wavelength. However, the heating of the bulk of the dust is inefficient. The reason is that once the UV photons get absorbed and re-emit as FIR photons, the chance of them being

<sup>8</sup>The median sSFRs at  $z = 2, 3, 4$ , and  $6$  of the MASSIVEFIRE sample are  $2.1 \times 10^{-10}$ ,  $5.8 \times 10^{-10}$ ,  $8.7 \times 10^{-10}$ , and  $3.3 \times 10^{-9} \text{ yr}^{-1}$ , respectively. SFRs are averaged over the past 20 Myr.

absorbed by dust again becomes much lower as a consequence of the declining opacity with wavelength ( $\kappa_\lambda \propto \lambda^{-2}$ ) (Scoville 2013). It is also interesting to note that both  $T_{\text{peak}}$  and  $T_{\text{mw}}$  are less correlated with SB when sSFR is averaged over longer period of time (Faucher-Giguère 2018; Feldmann 2017; Sparre et al. 2017). By averaging sSFR over a period of 100 Myr instead of 20 Myr, for example,  $\rho$  of the  $T_{\text{peak}}$  ( $T_{\text{mw}}$ ) versus SB relation declines from 0.55 (0.43) to 0.13 (0.22).

We note that comparing to the recent observations of the  $z = 2-4$  star-forming galaxies (Schreiber et al. 2015), the median sSFR of the MASSIVEFIRE galaxies is about 0.2–0.3 dex lower, but is still within the lower  $1\sigma$  limit of the observational data (see Feldmann et al. 2016; Feldmann et al. 2017). This discrepancy of sSFR is commonly seen in the current cosmological galaxy simulations. A systematic increase of SFR will lead to more heating to the ISM dust and hence higher simulated dust temperatures, but will not affect the slope of the sSFR versus  $T$  relation. The increment of  $T_{\text{peak}}$  is estimated to be about 1.7–2.6 K according to the sSFR versus  $T$  relation (Fig. 9), which appears to be insignificant compared to the scatter of the observational data (Fig. 4). It should also be noted that the impact of an increased sSFR on  $T_{\text{peak}}$  can easily be offset by an increase of dust mass, which can potentially be driven by an increased  $\delta_{\text{dzt}}$ , dust opacity, or gas metallicities – all these properties are currently uncertain at high redshifts.

Finally, we show the relation between dust temperatures and  $M_{\text{star}}$  in the *right-hand panels*. Looking at the upper panel, it is clear that  $T_{\text{peak}}$  has very weak correlation with  $M_{\text{star}}$ . This again shows that  $T_{\text{peak}}$  is strongly influenced by the emission from the warm dust that is associated with the recently formed young stars and does not have as strong correlation with the total stellar mass of a galaxy. In contrast,  $T_{\text{mw}}$  is less sensitive to the variance of recent star-forming conditions and therefore shows relatively small scatter at given  $M_{\text{star}}$  at each redshift. The normalization of the  $T_{\text{mw}}$  versus  $M_{\text{star}}$  relation increases with redshift, which is driven by the rise of  $\text{SFR}/M_{\text{dust}}$  (i.e. energy injection rate per unit dust mass). We also notice a slight increase of  $T_{\text{mw}}$  with  $M_{\text{star}}$ . This is owing to the decrease of  $M_{\text{star}}/M_{\text{dust}}$  with  $M_{\text{star}}$  of the MASSIVEFIRE sample. As a result,  $\text{SFR}/M_{\text{dust}}$  slightly increases with  $M_{\text{star}}$  [i.e.  $\text{SFR}/M_{\text{dust}} \propto \text{sSFR}(M_{\text{star}}/M_{\text{dust}}) \propto M_{\text{star}}^{0.3}$ ] at given redshift.

#### 4 (SUB)MILLIMETRE BROAD-BAND FLUXES

A major problem for probing the dust properties in the high-redshift ( $z > 4$ ) is that most observations of dust emission at such high redshift are limited to a single broad-band flux detected by ALMA band 7 or 6. Deriving infrared luminosities and hence SFRs of these  $z \gtrsim 4$  objects is very challenging without FIR constraints and depends highly on the assumed *equivalent* dust temperature for the flux-to-luminosity conversion. The same problem also applies to many faint (i.e. below a few mJy) submm-selected objects at lower redshift ( $2 < z < 4$ ) that do not have *Herschel* FIR coverage. Therefore, an accurate estimate of  $T_{\text{eqv}}$  of the adopted SED function for different redshifts is critical.

In this section, we will analyse the  $T_{\text{eqv}}$  distribution of galaxies at  $z = 2-6$  with the help of the MASSIVEFIRE sample. Specifically, in Section 4.1, we will examine the redshift evolution of  $T_{\text{eqv}}$  and its dependence on  $\delta_{\text{dzt}}$ , offering a ‘cookbook’ for converting between (sub)mm and  $L_{\text{IR}}$  observations. In Section 4.2, we will compare  $T_{\text{eqv}}$  with  $T_{\text{mw}}$  and  $T_{\text{peak}}$ , and provide a physical interpretation of this dust temperature.

#### 4.1 The flux-to-luminosity conversion

$L_{\text{IR}}$  is often extrapolated from a single broad-band (sub)mm flux given the lack of additional submm or FIR constraints. A typical approach is to assume that the SED has an OT-MBB (or G-MBB) shape with a chosen value of the dust temperature parameter. However, as we have shown in Fig. 3 and discussed in Section 3.1, choosing a dust temperature parameter that is not compatible with the assumed functional shape of SED can result in significant biases for estimating  $L_{\text{IR}}$  of a galaxy. By definition, this problem is avoided if the adopted dust temperature is chosen to be  $T_{\text{eqv}}$ .

With the MASSIVEFIRE sample, we are able to predict the full dust SED for the high-redshift ( $z = 2-6$ ) objects covering over two orders of magnitude of IR luminosity ( $L_{\text{IR}} \approx 10^{10}-10^{12} L_{\odot}$ ). We predict the observed flux densities at ALMA band 7 ( $S_{870\mu\text{m}}$ ) and band 6 ( $S_{1.2\text{mm}}$ ) given the SED and redshift as well as  $L_{\text{IR}}$ . Many of these objects have  $S_{870\mu\text{m}}$  ( $S_{1.2\text{mm}}$ )  $\gtrsim 0.1$  mJy, which are over the  $3\sigma$  detection limit of ALMA band 6 and 7 using a typical integration time of 1 h. With the calculated  $S_{870\mu\text{m}}$  (and  $S_{1.2\text{mm}}$ ) of each galaxy, we find the OT-MBB (with  $\beta = 2.0$ ) and GP-MBB functions (with  $\beta = 2.0$ ,  $\lambda_1 = 100 \mu\text{m}$ ,  $\alpha = 2.5$  and the suggested value of  $N_{\text{pl}}$  by Casey 2012), normalized to match their observed flux densities at both ALMA bands, that can predict their true  $L_{\text{IR}}$ . By adjusting the temperature parameter in the fitting function to match both observed submm flux density and true  $L_{\text{IR}}$ , we obtain  $T_{\text{eqv}}$ , i.e. *the value of  $T$  necessary for obtaining an accurate estimate of  $L_{\text{IR}}$  from the measured (sub)mm flux densities for each galaxy*.

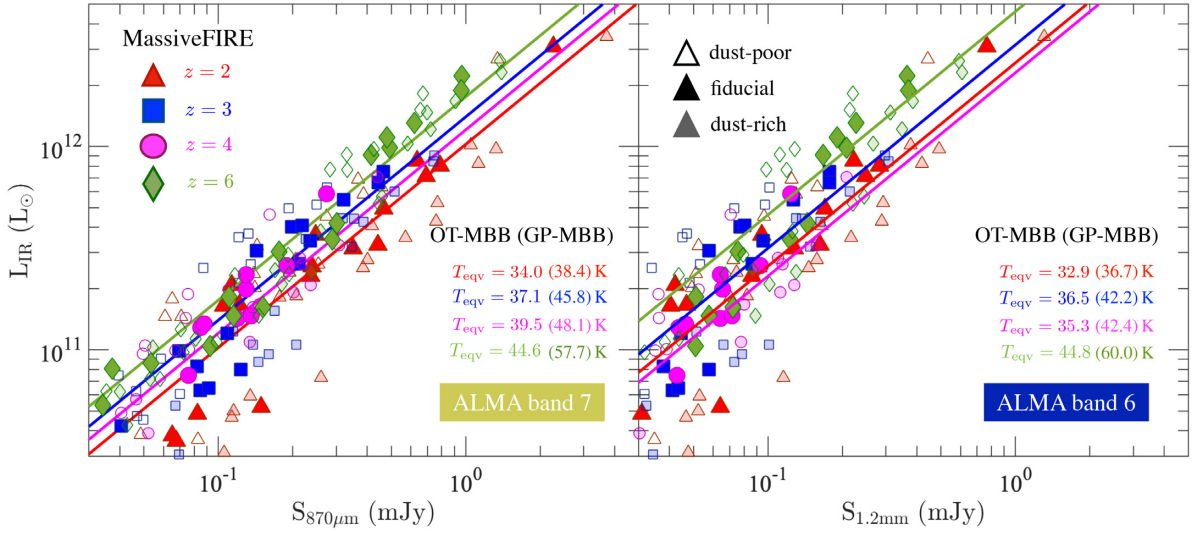
In Fig. 10, we show the relation of  $L_{\text{IR}}$  against  $S_{870\mu\text{m}}$  (*left-hand panel*) and  $S_{1.2\text{mm}}$  (*right-hand panel*) for the  $z = 2-6$  MASSIVEFIRE galaxies. For each redshift, we also show the expected  $L_{\text{IR}}$  versus  $S_{870\mu\text{m}}$  (and  $S_{1.2\text{mm}}$ ) relation using the mean  $T_{\text{eqv}}$  for galaxies above 0.1 mJy. The latter temperature is provided for the two different ALMA bands and for redshifts  $z = 2-6$ . We present the results for OT-MBB and GP-MBB functional shapes.

There appears to be a clear trend of increasing  $T_{\text{eqv}}$  with redshift, with either forms of fitting function (GP or OT-MBB) and with either ALMA band 6 or 7. This shows that a higher  $T_{\text{eqv}}$  is typically needed for deriving  $L_{\text{IR}}$  of galaxies at higher redshift. Using OT-MBB function, for example, the mean  $T_{\text{eqv}}$  increases from 34.0 K at  $z = 2$  (red triangles) to 44.6 K at  $z = 6$  (green diamonds) for ALMA band 7. Applying the typical  $T_{\text{eqv}}$  for  $z = 2$  to a  $z = 6$  galaxy will therefore lead to a significant underestimate of  $L_{\text{IR}}$ .

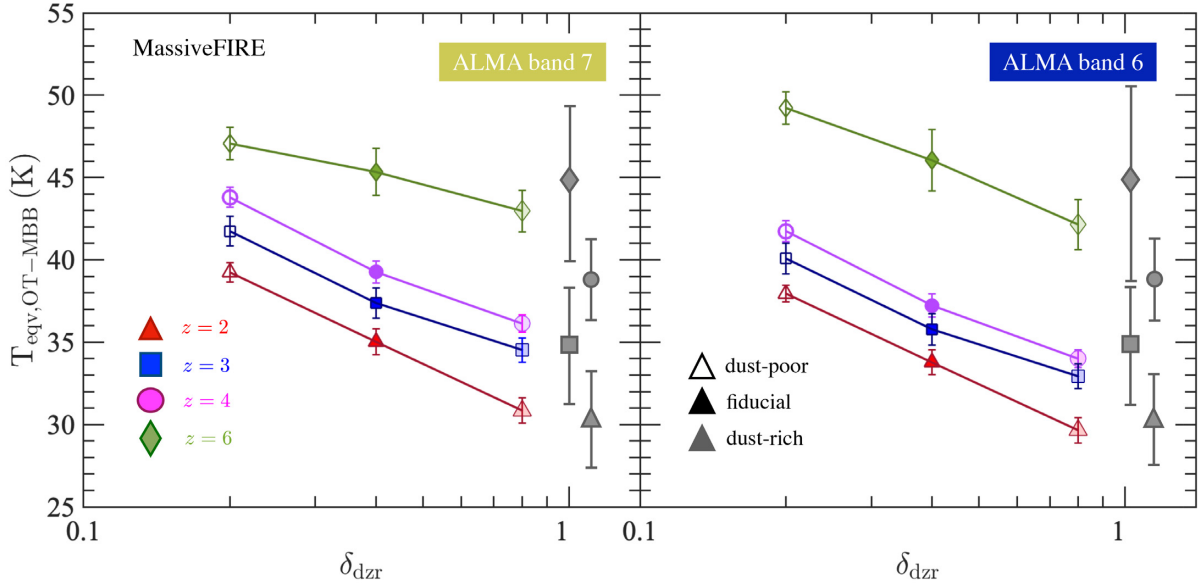
For the same redshift, the normalization of the  $L_{\text{IR}}$  versus  $S_{870\mu\text{m}}$  ( $S_{1.2\text{mm}}$ ) relation depends on dust mass. We explicitly show in Fig. 10 the result for dust-rich and dust-poor models. At fixed observed broad-band flux density, the  $L_{\text{IR}}$  of dust-rich galaxies lies systematically below the fiducial model (vice versa for dust-poor galaxies). This result indicates that a galaxy of given observed (sub)mm flux density tends to have lower (higher)  $L_{\text{IR}}$  if it contains more (less) amount of dust.

This finding can be understood as follows. By increasing the dust mass, both  $L_{\text{IR}}$  and  $S_{870\mu\text{m}}$  ( $S_{1.2\text{mm}}$ ) increase but the latter changes by a larger degree. Hence, the normalization of the relation declines. The increase of  $S_{870\mu\text{m}}$  ( $S_{1.2\text{mm}}$ ) is mainly driven by dust mass, as  $S_{870\mu\text{m}}$  ( $S_{1.2\text{mm}}$ ) is linearly scaled to  $M_{\text{dust}}$  (equation 11). On the other hand, the increase of  $L_{\text{IR}}$  is due to enhanced optical depth – a larger fraction of UV photons gets absorbed by dust and re-emitted in the infrared/submm. A lower  $T_{\text{eqv}}$  is therefore needed to account for the decrease of the normalization of the  $L_{\text{IR}}$  versus  $S_{870\mu\text{m}}$  ( $S_{1.2\text{mm}}$ ) relation with increasing dust mass. This anticorrelation of  $T_{\text{eqv}}$  with  $\delta_{\text{dzt}}$  is more clearly shown in Fig. 11.





**Figure 10.** Relation of  $L_{\text{IR}}$  versus  $S_{870\mu\text{m}}$  (left-hand panel) and  $S_{1.2\text{mm}}$  (right-hand panel) of our MASSIVEFIRE galaxy sample at  $z = 2-6$ . The unfilled, filled, and semitransparent symbols represent the result for a range of dust-to-metal ratios  $\delta_{\text{dzt}} = 0.2, 0.4$ , and  $0.8$ , respectively. The coloured lines show the  $L_{\text{IR}}$  versus  $S_{870\mu\text{m}}$  (and  $S_{1.2\text{mm}}$ ) relation expected from an OT-MBB function (equation 6, with fixed  $\beta = 2.0$ ) with an equivalent temperature ( $T_{\text{eqv}}$ , OT-MBB) that yields the  $L_{\text{IR}}$  of the MASSIVEFIRE sample at each redshift. The sample-average value of  $T_{\text{eqv}}$  for each redshift, ALMA band, and SED fitting function is labelled in the figure. Overall,  $T_{\text{eqv}}$  increases with redshift for the galaxies in our sample.



**Figure 11.** Relation of equivalent dust temperature ( $T_{\text{eqv}}$ ) versus dust-to-metal ratio ( $\delta_{\text{dzt}}$ ) of the  $z = 2-6$  MASSIVEFIRE sample.  $T_{\text{eqv}}$  is the effective dust temperature in the OT-MBB function (equation 6, with  $\beta = 2.0$ ) that yields the true  $L_{\text{IR}}$  of the galaxy from the flux densities at ALMA band 7 at  $870\mu\text{m}$  (left-hand panel) and band 6 at  $1.2\text{mm}$  (right-hand panel). The results of  $\delta_{\text{dzt}} = 0.4, 0.8$ , and  $0.2$  are shown with filled, semitransparent, and unfilled symbols, respectively. The grey error bars in both the panels represent the  $1\sigma$  dispersion of  $T_{\text{eqv}}$  for each redshift.  $T_{\text{eqv}}$  increases with redshift, and at the same redshift,  $T_{\text{eqv}}$  shows negative correlation with  $\delta_{\text{dzt}}$ .

We therefore provide a two-parameter fit for  $T_{\text{eqv}}$  with  $\delta_{\text{dzt}}$  and redshift as predictor variables. Using all the  $z = 2-6$  objects with  $S_{870\mu\text{m}} > 0.1\text{ mJy}$ , including the data for  $\delta_{\text{dzt}} = 0.2-0.8$ , we perform a multiple linear regression analysis

$$\log(T_{\text{eqv}}/25\text{ K}) = a + b \log(\delta_{\text{dzt}}/0.4) + c \log(1+z). \quad (15)$$

We present the best-fitting regression parameters  $a$ ,  $b$ , and  $c$  for ALMA band 6 and 7, and for OT-MBB and GP-MBB functions in Table 2. These derived scaling relations are useful for converting a

**Table 2.** Scaling relations between  $T_{\text{eqv}}$ ,  $\delta_{\text{dzt}}$ , and redshift.  $z = a + b \times x + c \times y$ , where  $z = \log(T_{\text{eqv}}/25\text{ K})$ ,  $x = \log(\delta_{\text{dzt}}/0.4)$ , and  $y = \log(1+z)$ .

	OT <sup>a</sup> (band 7)	OT <sup>a</sup> (band 6)	GP <sup>b</sup> (band 7)	GP <sup>b</sup> (band 6)
a	$-0.01 \pm 0.03$	$-0.05 \pm 0.04$	$0.00 \pm 0.05$	$-0.08 \pm 0.05$
b	$-0.13 \pm 0.03$	$-0.15 \pm 0.03$	$-0.22 \pm 0.04$	$-0.24 \pm 0.05$
c	$0.31 \pm 0.05$	$0.36 \pm 0.06$	$0.45 \pm 0.07$	$0.54 \pm 0.08$

<sup>a</sup>With fixed  $\beta = 2.0$ .

<sup>b</sup>With  $\lambda_1 = 100\mu\text{m}$ ,  $\beta = 2.0$ ,  $\alpha = 2.5$ , and the fiducial  $N_{\text{pl}}$  by Casey (2012).

measured (sub)mm flux density into  $L_{\text{IR}}$ , provided the redshift and dust-to-metal ratio of galaxy can be constrained.

The photometric redshift of the (sub)mm-detected galaxies can be determined when multiband optical and NIR data are available, and the more accurate spectroscopic redshift can subsequently be determined if several atomic/molecular emission lines (e.g. CO, C II, N II, O III) are identified (e.g. Laporte et al. 2017; Hashimoto et al. 2018; Patil et al. 2019). In contrast,  $\delta_{\text{dzt}}$  is more difficult to constrain from direct observation and is not yet well understood. Recent studies have reported differing results on how  $\delta_{\text{dzt}}$  depends on redshift and other galaxy properties (Inoue 2003; McKinnon, Torrey & Vogelsberger 2016; Wiseman et al. 2017; De Vis et al. 2019). We will discuss in more detail about the recent observations of  $\delta_{\text{dzt}}$  and the implication of the reduced  $\delta_{\text{dzt}}$  at high redshifts in Section 5.3.

#### 4.2 The equivalent dust temperature

$T_{\text{eqv}}$  depends on redshift and  $\delta_{\text{dzt}}$  in a clear and systematic manner, see Table 2. For the OT-MBB functional shape, for example,  $T_{\text{eqv}}$  scales as  $\propto (1+z)^{0.31} \delta_{\text{dzt}}^{-0.13}$  for ALMA band 7. This means that by applying a typical  $T_{\text{eqv}}$  for  $z = 2$  to a  $z = 6$  galaxy would lead to an underestimate of  $L_{\text{IR}}$  by a factor of  $\sim 4$  (equation 10). Also, at a given redshift, an order-of-magnitude increase of  $\delta_{\text{dzt}}$  corresponds to an  $\sim 0.13$  dex decrease of the best-fitting  $T_{\text{eqv}}$ . This corresponds to a decrease of  $L_{\text{IR}}$  by a factor of  $\sim 4$  (equation 10). Therefore, not taking the correlation of  $T_{\text{eqv}}$  with redshift and  $\delta_{\text{dzt}}$  into account can potentially lead to significant biases in the  $L_{\text{IR}}$  (and hence SFR) estimates.

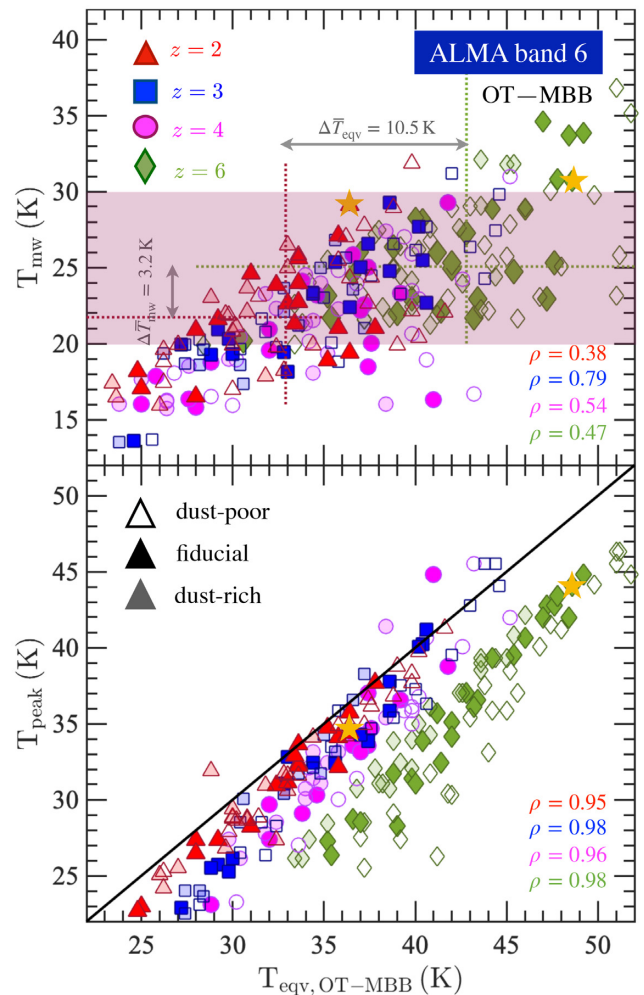
The scaling  $T_{\text{eqv}} \propto (1+z)^{0.31}$  (for band 7 and OT-MBB) is quantitatively similar to the one for  $T_{\text{peak}}$  (equation 12), meaning that  $T_{\text{eqv}}$  also evolves more quickly with redshift compared to  $T_{\text{mw}}$  (see the left-hand panels of Fig. 8). A natural question arises – what drives the evolution of  $T_{\text{eqv}}$  with redshift?

To answer this question, we show in Fig. 12 the  $T_{\text{eqv}}$  versus  $T_{\text{mw}}$  (upper panel) and  $T_{\text{eqv}}$  versus  $T_{\text{peak}}$  (lower panel) relations of the MASSIVEFIRE sample at  $z = 2-6$ . In this figure,  $T_{\text{eqv}}$  is calculated using an OT-MBB functional form (with fixed  $\beta = 2.0$ ) given a flux density at ALMA band 6. Using ALMA band 7 or a different form of MBB function results in qualitatively similar results and thus does not affect our conclusions.

It is clear from Fig. 12 that  $T_{\text{eqv}}$  is more strongly correlated with  $T_{\text{peak}}$  than  $T_{\text{mw}}$ , either by looking at the  $z = 2-6$  sample as a whole, or each individual redshift. For each redshift,  $T_{\text{peak}}$  scales approximately linearly with  $T_{\text{eqv}}$ , with a high Spearman correlation coefficient  $\rho \gtrsim 0.95$ . In contrast, the relation between  $T_{\text{mw}}$  and  $T_{\text{eqv}}$  is sublinear and shows large scatter. As shown in the upper panel, galaxies with similar  $T_{\text{mw}}$  can have very different  $T_{\text{eqv}}$  ( $\Delta T_{\text{eqv}} > 10$  K) and thus a large range of  $L_{\text{IR}}/S$  ratios (equation 10).

To understand the origin of the scatter in  $T_{\text{eqv}}$  and fixed  $T_{\text{mw}}$ , we selected two galaxies from the MASSIVEFIRE sample with similar  $T_{\text{mw}}$  ( $\approx 30$  K), one from  $z = 2$  and the other from  $z = 6$ , and study their SEDs and their  $T_{\text{eqv}}$  in more detail. The two galaxies are marked in both the panels of Fig. 12 by yellow asterisks, and their SEDs are presented in Fig. 3. The  $z = 6$  galaxy has  $T_{\text{eqv, OT-MBB}} = 49.1$  K, which is about 14 K higher than the  $z = 2$  galaxy.

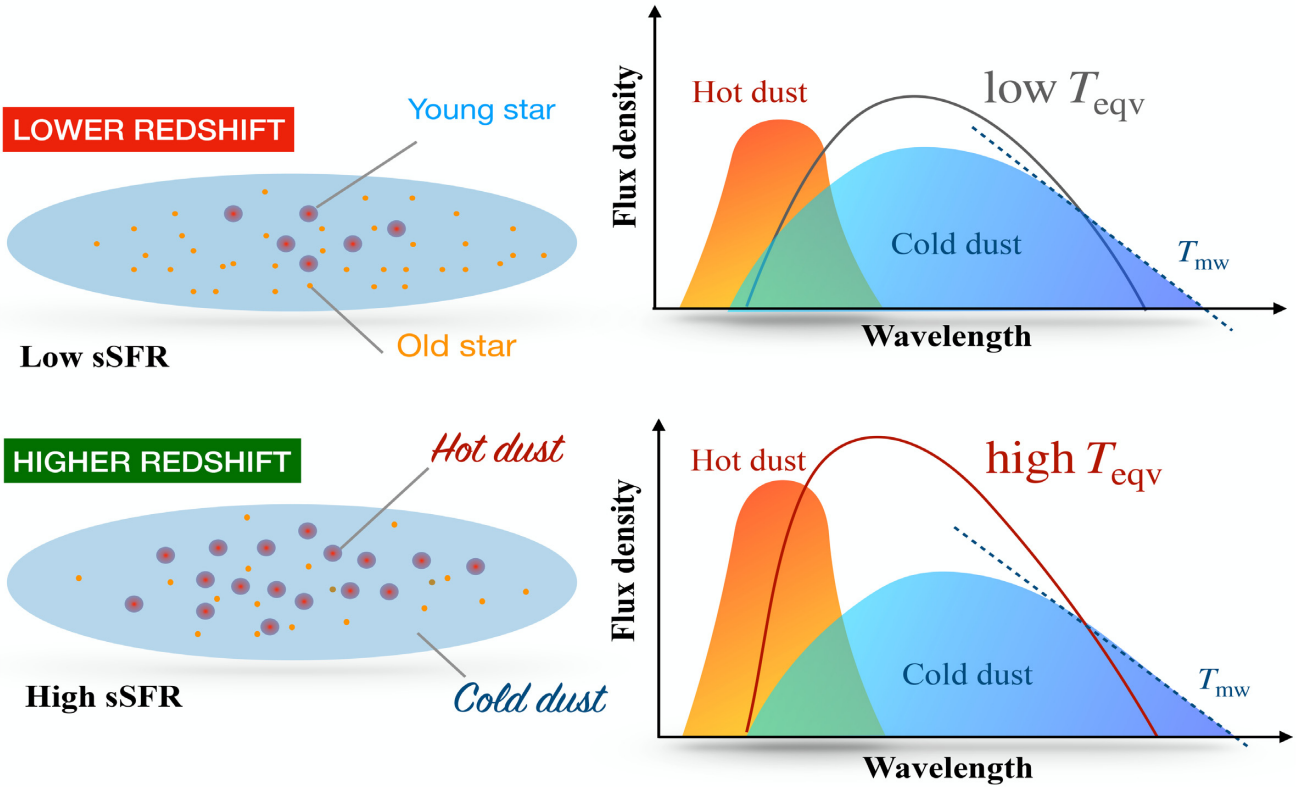
Fig. 3 shows that the two galaxies have different SED shapes at short wavelengths. The  $z = 6$  galaxy shows more prominent MIR emission due to its more active recent star formation. Its sSFR ( $= 5.0 \times 10^{-9} \text{ yr}^{-1}$ ) is about one order of magnitude higher than that of the  $z = 2$  galaxy. Young star clusters in this high-redshift galaxy efficiently heat the dense, surrounding dust, which boosts the



**Figure 12.** Relation of  $T_{\text{mw}}$  (upper panel) versus  $T_{\text{eqv, OT-MBB}}$  and  $T_{\text{peak}}$  (lower panel) versus  $T_{\text{eqv, OT-MBB}}$  of the  $z = 2-6$  MASSIVEFIRE galaxies, where  $T_{\text{eqv, OT-MBB}}$  is the equivalent dust temperature for the adopted OT-MBB function (equation 6, with fixed  $\beta = 2.0$ ) that yields the right  $L_{\text{IR}}$  from  $S_{1.2\text{mm}}$ . In the upper panel, the two horizontal dotted lines mark the median  $T_{\text{mw}}$  of the  $z = 2$  (red) and  $z = 6$  (green) samples, while the two vertical dotted lines mark their mean  $T_{\text{eqv, OT-MBB}}$ . The purple shaded box shows  $T_{\text{mw}} = 25 \pm 5$  K, where  $T_{\text{mw}} = 25$  K is the suggested dust temperature for estimating dust/gas mass using the RJ approach by Scoville et al. (2016) (Section 3.4.1). The two yellow asterisks in each panel mark the selected  $z = 2$  (left) and  $z = 6$  (right) galaxies. Their SEDs are shown in Fig. 3. The two galaxies have similar  $T_{\text{mw}}$ , but very different  $T_{\text{peak}}$  and  $T_{\text{eqv}}$ .  $T_{\text{eqv}}$  exhibits stronger correlation with  $T_{\text{peak}}$  than  $T_{\text{mw}}$ .

MIR emission and thus leads to a relatively high  $T_{\text{peak}}$  ( $= 44.6$  K) to account for the more prominent MIR emission of this galaxy. Furthermore, the  $z = 6$  galaxy is less dust-enriched than the  $z = 2$  galaxy (having only 1/7 of dust mass), and its  $\text{SFR}/M_{\text{dust}}$  ratio is roughly 4 times higher.

The increased  $\text{SFR}/M_{\text{dust}}$  ratio would leave an imprint on the temperature of the diffuse dust if the heat budget of the young stars were evenly distributed in the ISM dust. However, the bulk of the diffuse cold dust is clearly not heated efficiently as the two galaxies have almost the same  $T_{\text{mw}}$  (29.1 K versus 30.7 K). A number of factors can influence how efficiently the bulk of the dust is heated, such as the spatial distribution of dust in galaxy and the optical depth in vicinity of the star-forming cores (cf. Narayanan et al.



**Figure 13.** Schematic figure for the ‘two-phase’ model of ISM dust and the implication on the dust SED. Higher redshift galaxies have higher sSFR and more young ( $t_{\text{age}} \lesssim 10$  Myr) star clusters efficiently heat the dense dust in vicinity of the star-forming regions to high temperature. This hot dust component boosts the overall SED of galaxy at MIR. A higher equivalent temperature ( $T_{\text{eqv}}$ ) is thus needed to account for the more prominent MIR emission of galaxies at higher redshift.  $T_{\text{eqv}}$  is not well correlated with the mass-weighted temperature ( $T_{\text{mw}}$ ) of galaxy.  $T_{\text{mw}}$  is determined by the cold dust component and it sets the slope of the RJ tail.

2018b; Katz et al. 2019). These conditions can be significantly different among galaxies and therefore  $T_{\text{mw}}$  is not expected to be well correlated with  $T_{\text{peak}}$  (see the *upper right panel* of Fig. 8). This example strongly indicates that a ‘two-phase’ picture of ISM dust is needed to account for the discrepancy between  $T_{\text{mw}}$  and  $T_{\text{peak}}$  (see Fig. 13).

Clearly,  $T_{\text{eqv}}$  depends on the exact form of the MBB function and the observing frequency band. As is shown in Table 2,  $T_{\text{eqv}}$  is higher at  $z = 2$  by 0.07 dex and it increases faster with redshift at  $z = 2-6$  when a GP-MBB function is used. Using the same MBB function,  $T_{\text{eqv}}$  also appears to be slightly higher (by  $\sim 0.05$  dex at  $z = 2-4$  for an OT-MBB function) when a flux density is measured at ALMA band 7 than band 6. As  $T_{\text{eqv}}$  depends both on the specific form of MBB function and the observing wavelength,  $T_{\text{eqv}}$  should *not* be interpreted as a physical temperature but rather understood as a parametrization of SED shape.

It may appear reasonable to use sSFR as a predictor variable instead of  $(1+z)$ , given that the former depends strongly on redshift (Fig. 9) and is physically linked to the amount of hot dust in galaxies. The reasons for adopting  $(1+z)$  are twofold. First of all, observationally, redshift of the (sub)mm-selected galaxies can be accurately determined through atomic/molecular emission lines, as discussed in Section 4.1. sSFR estimates, however, are uncertain because SFRs derived based on the non- $L_{\text{IR}}$  indicators (e.g. UV continuum and  $H_{\alpha}$  flux) are uncertain due to the variation of the dust attenuation laws (Wilkins et al. 2012; Conroy 2013; Narayanan et al. 2018b). Secondly, the mapping between observed (sub)mm flux and rest-frame SED introduces an explicit redshift

dependence on  $T_{\text{eqv}}$  – as is shown in the lower panel of Fig. 12, the normalization of the  $T_{\text{peak}}$  versus  $T_{\text{eqv}}$  relation declines with redshift using the same functional form and the observing frequency, indicating that a higher  $T_{\text{eqv}}$  (i.e. a steeper MBB function) is needed to derive  $L_{\text{IR}}$  when the rest-frame observing wavelength gets closer to the emission peak. Therefore, the  $(1+z)$  term in equation (15) accounts both (indirectly) for the cosmic time dependence of the sSFR and (directly) for the redshift of electromagnetic radiation.

Finally, adding  $M_{\text{star}}$  as a predictor variable results in a regression coefficient for the  $M_{\text{star}}$  term being consistent with zero. This means that our obtained fitting functions for  $T_{\text{eqv}}$  do not depend on the selection function of  $M_{\text{star}}$  of the MASSIVEFIRE sample, which can be different from that of the observations. Replacing the dependence on  $\delta_{\text{drr}}$  by a dependence on  $M_{\text{star}}$  or  $Z_{\text{gas}}$ <sup>9</sup> leads to a decreased goodness of fit for  $T_{\text{eqv}}$ .

## 5 DISCUSSION

### 5.1 Deriving $M_{\text{dust}}$

Many dust-enshrouded galaxies at high redshift ( $z > 2$ ) have been detected at (sub)mm wavelengths in the past years, thanks to

<sup>9</sup> $Z_{\text{gas}}$  is calculated using gas particles with temperature between 7000 and 15000 K and density above  $0.5 \text{ cm}^{-3}$ , which represent the nebular gas where the strong nebular emission lines originate (Ma et al. 2016).



the unprecedented sensitivity of ALMA. These (sub)mm-detected objects often lack a reliable measure of FIR photometry and many are extremely faint at UV/optical wavelengths (e.g. Daddi et al. 2009; Walter et al. 2012; Riguccini et al. 2015; Franco et al. 2018). A reliable estimate of their dust mass from full SED fitting is often not possible (cf. Behrens et al. 2018).

In the optically thin regime, the flux density in the RJ tail has a simple analytic form (equation 11), and  $M_{\text{dust}}$  can be derived from the flux density given  $T_{\text{mw}}$  (Section 3.4.1). However, it is difficult to constrain  $T_{\text{mw}}$  of high-redshift galaxies when individual star-forming regions are not resolved.

Fortunately, we find that  $T_{\text{mw}}$  does not strongly vary from galaxy to galaxy. This is noteworthy, given that our sample spans a wide range of cosmic time ( $z = 2-6$ ), stellar mass ( $M_{\text{star}} = 10^9-10^{12} M_{\odot}$ ), sSFR ( $10^{-10}-10^{-8} \text{ yr}^{-1}$ ), and IR luminosities ( $L_{\text{IR}} = 10^9-3 \times 10^{12} L_{\odot}$ ). In particular, 68 per cent (i.e.  $1\sigma$ ) of the galaxies in our sample have mass-weighted dust temperatures  $T_{\text{mw}} = 25 \pm 5 \text{ K}$ , corresponding to a 20 per cent uncertainty of estimating the dust mass as the mass estimates scale only linearly with  $T_{\text{mw}}$ , while 90 per cent of our sample lies within  $T_{\text{mw}} = 25 \pm 8 \text{ K}$  (32 per cent uncertainty of  $M_{\text{dust}}$ ). Our findings support the empirical approach of adopting a constant  $T_{\text{mw}} = 25 \text{ K}$  to estimate the ISM mass of high-redshift galaxies via equation (11) and  $\delta_{\text{dgr}}$  (Scoville et al. 2016).

While adopting a constant  $T_{\text{mw}}$  is a good assumption to first order, and the only option if the (sub)mm flux density is measured at only a single wavelength, additional constraints on the SED may help to determine  $T_{\text{mw}}$  and improve the accuracy of measuring ISM masses. Specifically, in Section 3.4.3, we show that  $T_{\text{mw}}$  is well correlated with  $L_{\text{IR}}$  and that the redshift evolution of the  $L_{\text{IR}}$  versus  $T_{\text{mw}}$  relation is driven by the evolving dust mass. In fact,  $L_{\text{IR}}$ ,  $M_{\text{dust}}$ , and  $T_{\text{mw}}$  follow a tight scaling relation (equation 13) for  $T_{\text{mw}} \gg T_{\text{CMB}}$ . Hence, given  $S \propto M_{\text{dust}} T_{\text{mw}}$ , it should be possible to simultaneously infer  $M_{\text{dust}}$  and  $T_{\text{mw}}$  from a combined measurement of  $S$  and  $L_{\text{IR}}$ .

Recent studies have shown that the broad-band rest-frame  $8 \mu\text{m}$  luminosity,  $L_8$ , can be a rough tracer of  $L_{\text{IR}}$  over a range of galaxies (see Elbaz et al. 2011; Magdis et al. 2013; Murata et al. 2014; Schreiber et al. 2018). One important practical interest for using  $L_8$  is that it will be easily accessible by the upcoming James Webb Space Telescope (JWST) for galaxies up to  $z \sim 3$ . The unprecedented sensitivity of the Mid-Infrared Instrument (MIRI) onboard the JWST, covering the wavelength range of  $5-28 \mu\text{m}$ , will significantly enlarge the sample size of distant galaxies with measured MIR broad-band spectroscopy. We thus propose to use  $L_8$  to infer  $L_{\text{IR}}$  for the (sub)mm-detected galaxies at  $z \lesssim 3$  that have no constraint on SED shape near the emission peak (cf. Chary & Elbaz 2001; Elbaz et al. 2002; Reddy et al. 2006; Rieke et al. 2009; Santini et al. 2009; Reddy et al. 2010; Rujopakarn et al. 2013; Shipley et al. 2016; Azadi et al. 2018).

Hence, we propose to derive  $M_{\text{dust}}$  (as well as  $T_{\text{mw}}$ ) of high-redshift galaxies by combining MIR (e.g. from JWST) and FIR/submm (e.g. ALMA) data sets. Specifically, by combining equations (11) and (13), we obtain

$$\log \left( \frac{M_{\text{dust}}}{M_{\odot}} \right) = 1.23 \log \left( \frac{S}{\text{mJy}} \right) - 0.23 \log \left( \frac{L_{\text{IR}}}{L_{\odot}} \right) + \mathcal{F}(z)$$

$$\text{or } M_{\text{dust}} \propto \left( \frac{S}{L_{\text{IR}}} \right)^{0.23} S, \quad (16)$$

where  $\mathcal{F}(z) = -0.85 + 1.23 \log(\psi(z)\Gamma_{\text{RJ}})$  and  $\psi(z)$  has the unit of  $\text{mJy } M_{\odot}^{-1} \text{ K}^{-1}$ . Assuming that  $L_{\text{IR}} = \alpha L_8$  (Magdis et al. 2013;

Schreiber et al. 2018), we can rewrite the above equation as

$$\log \left( \frac{M_{\text{dust}}}{M_{\odot}} \right) = 1.23 \log \left( \frac{S}{\text{mJy}} \right) - 0.23 \log \left( \frac{L_8}{L_{\odot}} \right) + \mathcal{G}(z), \quad (17)$$

where  $\mathcal{G}(z) = -0.23 \log \alpha + \mathcal{F}(z)$ . In general,  $\Gamma_{\text{RJ}}$  is a function of  $T$  and equations (11) and (17) need to be solved numerically.

It is important to note that  $\alpha$  can depend on the variation in the detailed conditions of the star-forming regions. Recent observational evidence has shown that scatter in  $\alpha$  can be driven by certain galaxy properties, such as sSFR,  $Z_{\text{gas}}$ , and compactness of IR-emitting regions (e.g. Nordon et al. 2012; Elbaz et al. 2018; Schreiber et al. 2018). These dependences on intrinsic galaxy properties can then translate to an apparent dependence of  $\alpha$  on redshift and starburstiness (SB, see Section 3.4.4 for definition) of galaxy. Therefore, to improve the accuracy of  $L_{\text{IR}}$  estimates from  $L_8$ , one needs to rely on either a direct measurement or an observational proxy of these properties and/or galaxy redshift/SB. Furthermore, one should also caution the contribution of PAH molecules and AGN activity to the  $8 \mu\text{m}$  features (e.g. Siebenmorgen, Krügel & Spoon 2004; Pope et al. 2008; Kirkpatrick et al. 2012; Magdis et al. 2013; Stierwalt et al. 2014; Kirkpatrick et al. 2015; Roebuck et al. 2016; Lambrides et al. 2019). This topic is beyond the scope of this paper and we leave it to a future study.

According to equation (16), a factor of 2 uncertainty in  $L_{\text{IR}}$  translates into  $\sim 20$  per cent uncertainty in the derived dust mass, i.e. matches the intrinsic level of error of the constant  $T_{\text{mw}} = 25 \text{ K}$  method (Scoville et al. 2016). Therefore, increasing complexity by deriving  $T_{\text{mw}}$  and  $M_{\text{dust}}$  from  $L_{\text{IR}}$  and  $S$  will only be beneficial if  $L_{\text{IR}}$  can be constrained to within a factor of 2 or better.

Finally, we note that our sample does not include the most luminous submm galaxies that can have  $S_{870 \mu\text{m}}$  fluxes of much higher than a few mJy (e.g. Oteo et al. 2017; Oteo et al. 2018) and, hence, we cannot rule out that  $T_{\text{mw}}$  significantly exceeds 20–30 K in such objects. While submm-luminous galaxies are typically interpreted as having high SFRs, equation (11) shows that submm fluxes are simply the product of  $M_{\text{dust}}$  and  $T_{\text{mw}}$  ( $\Gamma_{\text{RJ}}$  is a weak function of  $T$ ). Hence, as long as  $T_{\text{mw}}$  is not significantly higher in these objects, a straightforward interpretation is that the most submm-luminous galaxies are those with the highest  $M_{\text{dust}}$ . In fact,  $S_{870 \mu\text{m}}$  ( $S_{1.2 \text{ mm}}$ ) is nearly doubled as  $M_{\text{dust}}$  gets doubled (by comparing the dust-rich and fiducial cases in Fig. 10), while  $T_{\text{mw}}$  decreases by only  $\sim 1 \text{ K}$  (i.e.  $\lesssim 5$  per cent). *This example also suggests that caution needs to be taken when directly converting (sub)mm flux densities to  $L_{\text{IR}}$  (and SFR), without taking into account how dust mass (and optical depth) alters the SED shape of galaxy* (cf. Scoville & Kwan 1976; Hayward et al. 2011; Scoville 2013; Safarzadeh et al. 2016, and see the lower panel of Fig. 1, where we show how SED is altered by  $M_{\text{dust}}$  without having a different SFR of galaxy).

## 5.2 The increase of $T_{\text{eqv}}$ with redshift and its observational evidence

Adopting  $T_{\text{eqv}}$  and an SED shape is another way to estimate the IR luminosity from submm fluxes, see Section 4.2. Hence, if  $T_{\text{eqv}}$  is known, it is possible to use the approach described in the previous section to infer dust masses and mass-weighted temperatures. This could be a particularly useful approach at  $z \gtrsim 3$ , where the potential MIR diagnostics redshift out of the wavelengths accessible by JWST.

Using the simplified functional forms of SED (an OT-MBB or GP-MBB), the obtained  $T_{\text{eqv}}$  increases monotonically with redshift

(Table 2). The typical  $T_{\text{eqv}}$  of a  $z = 6$  galaxy is as high as 45–50 K for an OT-MBB function, which is significantly higher than the mean  $T_{\text{mw}}$  ( $\sim 25$  K) at this redshift. This result is consistent with what has been implied by some recent observational findings, including the unusual relationship between the IR excess ( $\text{IRX} \equiv L_{\text{IR}}/L_{\text{UV}}$ ) and the UV spectral slope ( $\beta$ ) of the Lyman break galaxies (LBGs) at high redshifts.

Empirically, IRX is used as a proxy of total dust mass and  $\beta$  is a measure of dust column density. It has been found that galaxies between  $0 < z < 4$  follow a well-defined sequence on the IRX– $\beta$  diagram, although there exists non-trivial scatter that depends on the stellar populations and the detailed dust and ISM properties (Meurer et al. 1995; Meurer, Heckman & Calzetti 1999; Siana et al. 2009; Takeuchi et al. 2012; Faisst et al. 2017; Popping, Puglisi & Norman 2017; Narayanan et al. 2018a; Ma et al. 2019). However, recent observations of higher redshift LBGs show evident IRX deficit (e.g. Capak et al. 2015; Bouwens et al. 2016). A significant fraction of the  $z > 4$  UV-selected galaxies have no observed dust continuum at ALMA bands. For the ALMA-detected objects, their estimated IRX appears to be significantly lower than the value inferred from their measured  $\beta$  using the canonical IRX– $\beta$  relations found by the local samples. The IRX deficit of the selected high-redshift LBGs is challenging to explain with the current dust attenuation models (Faisst et al. 2017; Ferrara et al. 2017; Narayanan et al. 2018a).

Instead of the high-redshift populations having significantly different dust properties, an alternative solution is that they have a higher  $T_{\text{eqv}}$ , which results in a higher derived  $L_{\text{IR}}$  with a given observed (sub)mm flux density. For example, Bouwens et al. (2016) report that among 330 LBGs in the Hubble Ultra Deep Field spanning the redshift range of  $z = 2$ –10, only 6 were detected at  $\gtrsim 2\sigma$  at 1.2mm (band 6) by the ALMA Spectroscopic Survey (ASPECS). This is significantly lower than the number (35) extrapolated from the  $z = 0$  IRX– $\beta$  relation based on their UV properties and by assuming a constant (equivalent) dust temperature of 35 K. The authors suggest that using a monotonic increase of (equivalent) dust temperature with redshift, i.e.  $T \propto (1+z)^{0.32}$  (an OT-MBB function is assumed), the number of detected sources can be consistent with the SMC IRX– $\beta$  relation, which is derived based on the local metal-poor populations (Siana et al. 2009). Encouragingly, this suggested redshift dependence of temperature well agrees with that of  $T_{\text{eqv}}$  found by the MASSIVEFIRE sample (see Table 2). This indicates that a significant deviation of the dust properties of the high-redshift UV-selected populations is not needed for explaining their observed IRX deficit (cf. Casey et al. 2018b).

Finally,  $T_{\text{eqv}}$  should not be deemed equivalent to the *mean intensity of the radiation field*,  $\langle U \rangle \propto L_{\text{IR}}/M_{\text{dust}}$  (e.g. Draine & Li 2007), despite that the latter has also been found to increase with redshift (e.g. Magdis et al. 2012; Béthermin et al. 2015; Magdis et al. 2017; Schreiber et al. 2018). This is because  $T_{\text{eqv}}$  is a parametrization of SED shape, and it depends on both the assumed functional form of SED as well as the observing wavelength (Section 4.2), while  $\langle U \rangle \propto T_{\text{mw}}^{5.4}$ , equation 13 represents the physical dust temperature. It is also important to note that since  $T_{\text{mw}}$  not only evolves with redshift, but also shows a clear dependence on  $L_{\text{IR}}$  at fixed redshift (see the *lower left panel* of Fig. 8), the observed redshift evolution of  $\langle U \rangle$  therefore can potentially depend on the selection function of  $L_{\text{IR}}$ . *The selection bias caused by using a flux-limited sample (galaxies at higher redshift are confined to higher  $L_{\text{IR}}$ ) can lead to a steeper increase of  $\langle U \rangle$  with redshift than is measured at fixed  $L_{\text{IR}}$ .*

### 5.3 The dependence of $T_{\text{eqv}}$ on $\delta_{\text{dzt}}$

Fig. 11 shows that  $T_{\text{eqv}}$  is anticorrelated with  $\delta_{\text{dzt}}$  at fixed redshift. Specifically, an order-of-magnitude decrease in  $\delta_{\text{dzt}}$  translates to  $\sim 0.13$  dex increase of required  $T_{\text{eqv}}$ , corresponding to a factor of  $\sim 4$  increase of  $L_{\text{IR}}$  (equation 10). As noted before (Sections 4.1 and 5.1),  $\delta_{\text{dzt}}$  directly affects the total dust mass (and optical depth) of a galaxy, thereby altering its SED shape. Hence, it can be one source of uncertainty in estimating  $L_{\text{IR}}$  through  $T_{\text{eqv}}$ .

Observationally, while  $\delta_{\text{dzt}}$  has been found to be fairly constant across a wide range of galaxies at low redshifts by different studies (Issa, MacLaren & Wolfendale 1990; Lisenfeld & Ferrara 1998; James et al. 2002; Galliano et al. 2005; Watson 2011), there is also evidence of a reduced  $\delta_{\text{dzt}}$  in the low-metallicity environments (Herrera-Camus et al. 2012; Rémy-Ruyer et al. 2014; De Cia et al. 2016; McKinnon et al. 2016; De Vis et al. 2017; Chiang et al. 2018; De Vis et al. 2019). This can imply a lower  $\delta_{\text{dzt}}$  in higher redshift galaxies since they are known to have lower metallicities than the low-redshift galaxies (Erb et al. 2006; Finlator & Davé 2008; Maiolino et al. 2008; Mannucci et al. 2010; Davé, Finlator & Oppenheimer 2011; Lilly et al. 2013; Ma et al. 2016; Onodera et al. 2016). A direct implication of the decrease of  $\delta_{\text{dzt}}$  with redshift is that it can further mitigate the IRX deficit problem of the high-redshift LBGs (see Section 5.2). A reduced  $\delta_{\text{dzt}}$  leads to an additional increase of  $T_{\text{eqv}}$  and hence a higher upper confidence limit of the IRX of the undetected objects, meaning that the dust properties of these high-redshift LBGs are more probable to be consistent with the canonical dust attenuation laws that are derived based on the low-redshift observations.

However,  $\delta_{\text{dzt}}$  of high-redshift galaxies is not yet well understood. Recent studies based on foreground absorbers towards  $\gamma$ -ray burst (GRB) afterglows and quasars (QSOs) as well as distant lensed galaxies have shown different trends of how  $\delta_{\text{dzt}}$  depends on redshift (Dai & Kochanek 2009; Chen et al. 2013; De Cia et al. 2013; Zafar & Watson 2013; De Cia et al. 2016; Wiseman et al. 2017), which can be due to selection effects and observational uncertainties (Mattsson et al. 2014), as well as the different choice of  $\delta_{\text{dzt}}$  measures (Wiseman et al. 2017). Moreover, because these studies are often limited to only a handful of galaxies, it becomes difficult to distinguish the explicit redshift dependence of  $\delta_{\text{dzt}}$  from the intrinsic correlation with the galaxy properties, although some studies based on the QSO-damped Lyman- $\alpha$  absorbers (QSO-DLAs) find that the QSO-DLAs over a range of redshifts ( $z = 2$ –6) follow a similar  $\delta_{\text{dzt}}$ – $Z_{\text{gas}}$  relation (De Cia et al. 2013, 2016; Wiseman et al. 2017).

Observationally, gas-phase galaxy metallicities ( $Z_{\text{gas}}$ ) can be derived using the ratios between (rest-frame) optical auroral and nebular line fluxes and with calibration on theoretical models (e.g. Kewley & Dopita 2002; Kewley & Ellison 2008; Maiolino et al. 2008; Zahid, Kewley & Bresolin 2011; Steidel et al. 2014). However, this method can only be used for galaxies up to  $z \sim 3$ , above which the emission lines redshift out of the wavelengths of the current ground-based NIR spectrographs. An alternative method is to use the equivalent width of (rest-frame) UV absorption features, which has been used for galaxies up to  $z \sim 5$ , but is still limited because of the faintness of the features (Heckman et al. 1998; Eldridge & Stanway 2012; Faisst et al. 2016).

To overcome these shortcomings, Rigopoulou et al. (2018) have recently proposed a new method of using the (rest-frame) FIR  $[\text{OIII}]$  88  $\mu\text{m}/[\text{NII}]$  122  $\mu\text{m}$  line ratio for probing the gas metallicities of galaxies at  $z > 4$ , where both characteristic lines shift to the submm range that is accessible with ALMA. Using the previously reported FIR line measurements of a sample of local normal and star-forming

galaxies by the *Infrared Space Observatory* (ISO; Kessler et al. 1996), Rigopoulou et al. (2018) find that the derived galaxy mass–metallicity relation is consistent with the result derived using optical emission lines (Tremonti et al. 2004). The gas metallicities of three  $z = 2$ –3 submm-luminous galaxies derived using *Herschel* measurements are also in good agreement with the high-redshift relationships previously derived by Maiolino et al. (2008) and Mannucci et al. (2010), despite that the stellar mass estimates of these obscured dusty high-redshift galaxies have large uncertainties. These results suggest that FIR emission lines could be promising tool for estimating  $Z_{\text{gas}}$  of  $z > 4$  galaxies.

Therefore, if  $Z_{\text{gas}}$  could be used as a predictor for  $\delta_{\text{dzt}}$  (which is currently uncertain), it will further help improve the accuracy of  $T_{\text{eqv}}$  (and hence  $L_{\text{IR}}$ ) estimates at  $z > 4$ , which is important since the MIR diagnostics for  $L_{\text{IR}}$  are inaccessible by JWST at this epoch (Section 5.1). This will in turn help improve our constraints on the total obscured SFR in the early Universe, where currently only UV-based SFR estimates are available (Casey et al. 2018a,b).

#### 5.4 The subresolution structure of the birth-clouds

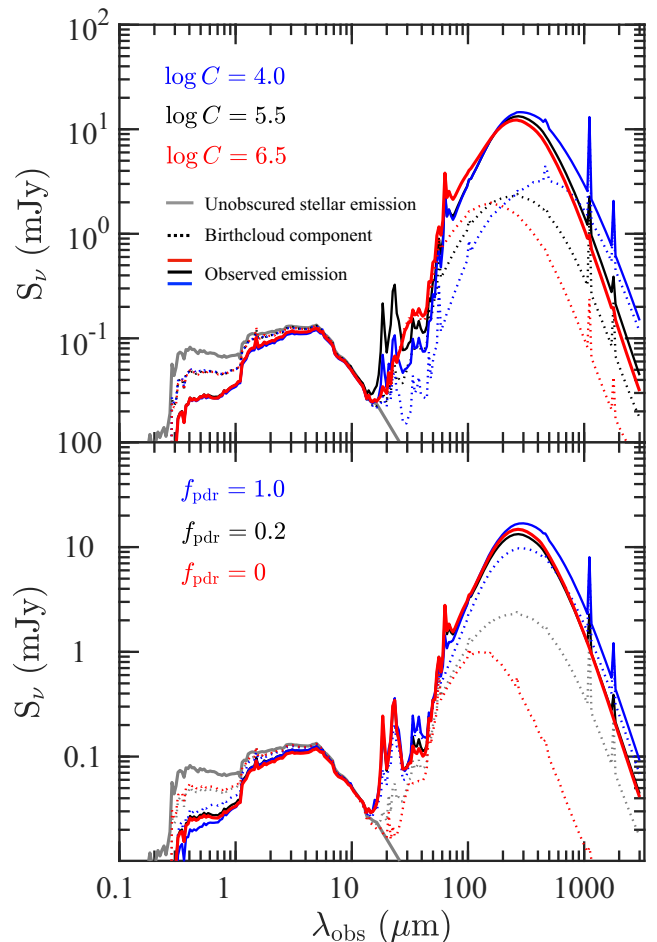
Observational evidence has indicated that young star clusters reside in dense dusty birth-clouds (e.g. Calzetti et al. 1997; Tuffs et al. 2004; Wild et al. 2011; Price et al. 2014; Koyama et al. 2015). To check the uncertainty arising from potentially unresolved small-scale ISM structure, we have repeated the analysis presented in this paper with additional RT analysis by SKIRT as Liang et al. (2018), where we include a subgrid model for birth-clouds embedding the young stars (our ‘alternative’ RT model). We summarize the details of this subgrid model and the main results from this model in this subsection.

In brief, all the young star particle of a galaxy that has formed less than 10 Myr ago is assigned a MAPPINGSIII source SED (Groves et al. 2008). MAPPINGSIII SED templates are parametrized by the SFR and the metallicity of the star-forming regions, the pressure of the ambient ISM, the H II region compactness ( $\log C$ ), and the covering fraction of the associated PDR ( $f_{\text{pdr}}$ ).

To explore how our results depend on this choice, the *upper* and *lower* panels in Fig. 14 show the overall SED of one of our galaxies for different values of  $\log C$  and  $f_{\text{pdr}}$ , respectively. As  $\log C$  increases, the birth-clouds become more compact and the dust associated with the clouds attains higher mean temperature because of the stronger incident radiation on to dust grains. The source SED of this dust component (shown with dashed lines) shifts to shorter wavelength, and so does the overall SED of the galaxy.  $f_{\text{pdr}}$  is a measure of the survival time-scale of birth-clouds (Jonsson et al. 2010). Increasing  $f_{\text{pdr}}$  results in a larger fraction of the stellar emission being absorbed by dust in the birth-clouds, which results in more energy being re-emitted as IR light. The mean dust temperature of the birth-clouds, however, decreases. For the total emission of galaxy, a higher  $f_{\text{pdr}}$  typically leads to higher  $L_{\text{IR}}$ . Whether the emission peak of the *overall* SED shifts to shorter or longer wavelengths with  $f_{\text{pdr}}$ , however, depends on the value of  $\log C$ .

The subgrid model has minor impact on  $T_{\text{mw}}$  of galaxies.  $T_{\text{mw}}$  increases with  $\log C$  at fixed  $f_{\text{pdr}}$ , and decreases with  $f_{\text{pdr}}$  at fixed  $\log C$ . The reason is that the photons emitted from the birth-clouds are more energetic if the birth-clouds are more compact (higher  $\log C$ ) and less dust-obscured (low  $f_{\text{pdr}}$ ). But the resulting difference of  $T_{\text{mw}}$  is typically no more than  $\pm 1$  K ( $\pm 5$  per cent) by exploring the parameter space of the MAPPINGSIII model.

$T_{\text{peak}}$ , however, is more sensitive to the uncertainty of the small-scale ISM structure. In some cases, especially for strongly star-

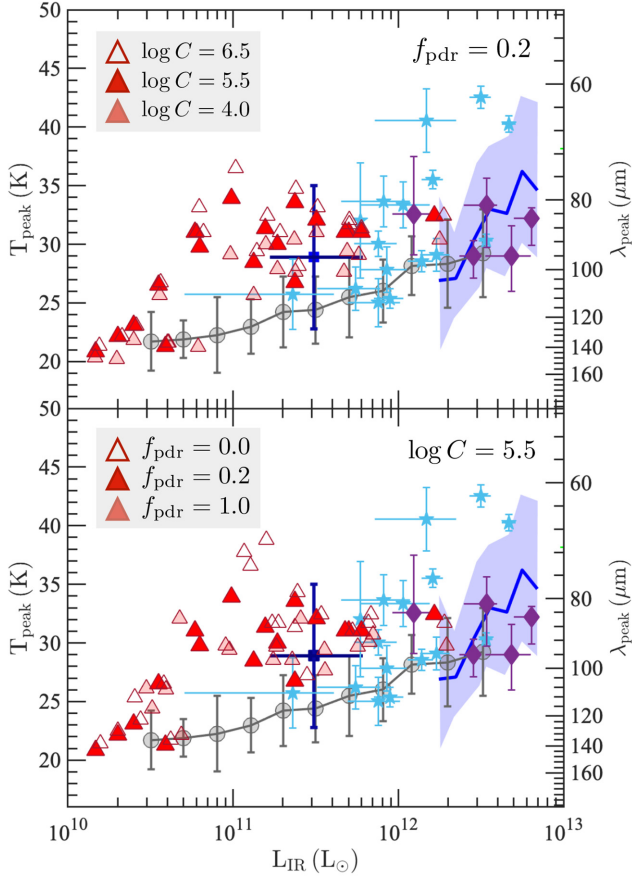


**Figure 14.** SEDs of a  $z = 2$  MASSIVEFIRE galaxy generated by different dust models. In the left-hand panel, we show the observed SEDs for  $\log C = 6.5$  (red), 5.5 (black), and 4.0 (blue) with fixed  $f_{\text{pdr}} (= 0.2)$ . In the right-hand panel, we show the result for  $f_{\text{pdr}} = 0$  (red),  $f_{\text{pdr}} = 0.2$  (black), and  $f_{\text{pdr}} = 1.0$  (blue) with fixed  $\log C (= 5.5)$ . In each panel, the grey curve shows the intrinsic stellar emission, while the solid red, black, and blue curves show the observed SEDs, each corresponding to a different dust model. Source SEDs from birth-clouds associated with the star-forming regions are shown with dotted lines with the corresponding colour for each model.

forming galaxies,  $T_{\text{peak}}$  can differ by much as 10 K when the MAPPINGSIII parameters are varied.  $T_{\text{peak}}$  is typically higher with increasing  $\log C$ , and for low/intermediate ( $\sim 4.0$ – $5.5$ ) value of  $\log C$ , decreases with  $f_{\text{pdr}}$ . For the  $z = 2$  MassiveFIRE sample,  $\log C = 6.5$  (max) leads to a median  $T_{\text{peak}}$  higher than  $\log C = 4.0$  (min) by about 4 K with fixed  $f_{\text{pdr}} = 0.2$ , and  $f_{\text{pdr}} = 1.0$  (max) yields a median  $T_{\text{peak}}$  lower than  $f_{\text{pdr}} = 0$  (min) by about 2.5 K with fixed  $\log C = 5.5$ . Uncertainty of the small-scale ISM conditions could introduce scatter in the observed  $T_{\text{peak}}$  versus  $L_{\text{IR}}$  relation in addition to galaxy-by-galaxy variations of  $\delta_{\text{dzt}}$  (Fig. 15).

Including the subgrid birth-cloud model strengthens the correlation between  $T_{\text{peak}}$  (and  $T_{\text{eqv}}$ ) and sSFRs of galaxies. By pre-processing starlight in birth-clouds, the range of the physical conditions surrounding star-forming regions is reduced. We note, however, that none of the trends reported in this paper change on a qualitative level by including or excluding the MAPPINGSIII birth-cloud model.





**Figure 15.** The same as the *upper panel* of Fig. 4 except that the data of the  $z = 2$  MASSIVEFIRE sample are produced by including the MAPPINGSIII birth-cloud model (Section 5.4). In the *upper panel*, we show the result for  $\log C = 6.5$  (unfilled),  $5.5$  (filled), and  $4.0$  (semitransparent) with fixed  $f_{\text{pdr}} (= 0.2)$ , while in the *lower panel*, we show the result for  $f_{\text{pdr}} = 0$  (unfilled),  $f_{\text{pdr}} = 0.2$  (filled), and  $f_{\text{pdr}} = 1.0$  (semitransparent) with fixed  $\log C (= 5.5)$ . For all different models,  $\delta_{\text{dzt}} = 0.4$ .

## 6 SUMMARY AND CONCLUSION

In this paper, we study dust temperatures of high-redshift galaxies and their scaling relationships with the help of cosmological zoom-in simulations and dust RT modelling. Our sample consists of massive ( $M_{\text{star}} > 10^{10} M_{\odot}$ )  $z = 2-6$  galaxies extracted from the MASSIVEFIRE suite (Feldmann et al. 2016; Feldmann et al. 2017), a set of cosmological hydrodynamic zoom-in simulations from the FIRE project (Hopkins et al. 2014). The sample encompasses 18 central galaxies at  $z = 2$  and their MMPs up to  $z = 6$ , together with a disjoint set of 11 central galaxies at  $z = 6$ . We generate FIR-to-mm broad-band fluxes and spectra for our galaxy sample with SKIRT.

We explicitly define and discuss four different dust temperatures that are commonly used in the literature,  $T_{\text{mw}}$ ,  $T_{\text{peak}}$ ,  $T_{\text{eff}}$ , and  $T_{\text{eqv}}$ .  $T_{\text{mw}}$  is the physical, mass-weighted temperature that can be extracted from RT analysis, but is often not easily accessible to observations.  $T_{\text{eff}}$  and  $T_{\text{peak}}$  are derived from SED fitting:  $T_{\text{eff}}$  is the  $T$  parameter in the best-fitting MBB function and  $T_{\text{peak}}$  is the inverse of emission peak wavelength. These two are the temperatures that are often adopted for analysing large statistical sample of galaxies by observational studies. And finally,  $T_{\text{eqv}}$  is the temperature one needs to convert single (sub)mm data to total IR luminosity based on an assumed SED shape.

The main findings of this paper are:

- (i) FIRE simulations together with RT processing successfully reproduce  $T_{\text{peak}}$  of  $z = 2$ ,  $L_{\text{IR}} \gtrsim 10^{11} L_{\odot}$  galaxies, in good agreement with recent observations (Fig. 4). The observational data show large scatter, which may be driven by galaxy-to-galaxy variations of  $\delta_{\text{dzt}}$  as well as local variations in the physical conditions of unresolved birth-clouds embedding young star clusters (Sections 3.3 and 5.4).
- (ii)  $T_{\text{mw}}$  is only weakly correlated with  $T_{\text{peak}}$  over  $z = 2-6$  (Fig. 8). The former sets the slope of the RJ tail (Fig. 3), and is the temperature needed for estimating dust and gas mass of distant galaxies (Fig. 6). Using  $T_{\text{peak}}$ , or  $T_{\text{eff}}$  (e.g. derived from full SED fitting), which is strongly correlated with  $T_{\text{peak}}$ , can lead to a systematic bias/error of the derived dust/gas mass, and may lead to an inaccurate interpretation of the star-forming conditions in high-redshift galaxies (Section 3.4.1).
- (iii)  $T_{\text{peak}}$  is well correlated with sSFR ( $\rho \sim 0.7$ ) (Fig. 8). Recently formed stars efficiently heat the dense, warm dust in the close vicinity of star-forming regions. The emission from this warm dust component boosts the overall dust SED at MIR, and helps to shift the emission peak to shorter wavelength (Fig. 13).  $T_{\text{mw}}$  is less well correlated with sSFR ( $\rho \sim 0.55$ ) and the scaling relation shows a flatter slope ( $\Delta T_{\text{mw}} \propto \text{SB}^{4.0}$  versus  $\Delta T_{\text{peak}} \propto \text{SB}^{8.7}$ ). The bulk of the cold diffuse dust is not as effectively heated as the warm dust component (Section 3.4.4).
- (iv)  $T_{\text{peak}}$  scales as  $(1+z)^{0.25}$  at fixed  $L_{\text{IR}}$  between  $z = 2$  and 6 driven by the increasing sSFR at higher redshift, which is consistent with recent observations (Section 3.4.3).  $T_{\text{mw}}$  evolves only weakly with redshift at fixed  $L_{\text{IR}}$  at  $z = 2-6$  (Fig. 8).
- (v) Of the galaxies in our sample, 68 per cent have mass-weighted dust temperatures  $T_{\text{mw}} = 25 \pm 5$  K (Fig. 12). This temperature range corresponds to an uncertainty of 20 per cent in estimating  $M_{\text{dust}}$  from a single submm band since the mass estimates scale only *linearly* with  $T_{\text{mw}}$ . Furthermore, 90 per cent of our sample lies within  $T_{\text{mw}} = 25 \pm 8$  K. Our findings support the empirical approach of adopting a constant  $T_{\text{mw}} = 25$  K to estimate the ISM mass of high-redshift galaxies (Scoville et al. 2016).
- (vi)  $T_{\text{mw}}$  is well correlated with  $L_{\text{IR}}$  at  $T_{\text{mw}} \gg T_{\text{CMB}}$  at a given redshift (Fig. 8). The normalization of this relation evolves weakly with redshift but the slope does not change. At higher redshift, galaxies of the same  $L_{\text{IR}}$  have higher  $T_{\text{mw}}$  but lower  $M_{\text{dust}}$ . Using the  $z = 2-6$  sample, we derive the scaling relation  $L_{\text{IR}} \propto M_{\text{dust}}^{1.0} T_{\text{mw}}^{5.4}$ , which appears to be shallower than the classical  $L_{\text{IR}} \propto M_{\text{dust}}^{T^{4+\beta}}$  ( $\beta \approx 2.0$  for our adopted dust model) relation expected from the optically thin assumption (Section 3.4.3).
- (vii) We propose to use this scaling relation to derive  $M_{\text{dust}}$  (and  $T_{\text{mw}}$ ) of high-redshift (sub)mm-detected galaxies, assuming that their  $L_{\text{IR}}$  can be constrained, for example, via the mid-IR luminosity probed by the *Spitzer* telescope and the upcoming JWST. We showed that this method improves over the  $T_{\text{mw}} = 25$  K approach if  $L_{\text{IR}}$  can be constrained to within a factor of 2 or better (Section 5.1).
- (viii)  $T_{\text{eqv}}$  increases with redshift, meaning that a higher temperature is needed to convert observed (sub)mm broad-band fluxes to  $L_{\text{IR}}$  (and hence SFRs) of galaxies at higher redshift.  $T_{\text{eqv}}$  is tightly correlated ( $\rho \gtrsim 0.95$ ) with  $T_{\text{peak}}$ , a much stronger correlation than with  $T_{\text{mw}}$  (Fig. 12). In particular, two galaxies at different redshifts can have very different  $T_{\text{eqv}}$  ( $\Delta T_{\text{eqv}} > 10$  K) but similar  $T_{\text{mw}}$  (Section 4).
- (ix) We find an anticorrelation between  $T_{\text{eqv}}$  and the dust-to-gas ratio,  $\delta_{\text{dzt}}$ . Hence, at a given redshift, dust-poorer galaxies need, on the average, a higher  $T_{\text{eqv}}$  for the (sub)mm-flux-to-IR-luminosity conversion. We express  $T_{\text{eqv}}$  as a power-law function of  $\delta_{\text{dzt}}$  and

( $1 + z$ ), and perform linear regression analysis using the MASSIVE-FIRE sample at  $z = 2-6$ . The best-fitting parameters of the scaling relation are provided in Table 2. We present the result for both ALMA band 6 and 7. We propose to apply the scaling relation to more accurately convert between (sub)mm flux and IR luminosity (and SFR) of high-redshift galaxies (Section 4).

To summarize our results, we find that the observationally derived temperatures, in particular,  $T_{\text{peak}}$ , generally differ from  $T_{\text{mw}}$ .  $T_{\text{peak}}$  shows a steeper slope and a stronger correlation with sSFR, and evolves more quickly with redshift compared with  $T_{\text{mw}}$ . We also find that  $T_{\text{eqv}}$  is more strongly correlated with  $T_{\text{peak}}$  than with  $T_{\text{mw}}$ .

The difference between  $T_{\text{peak}}$  and  $T_{\text{mw}}$  may be understood by a ‘two-phase’ picture of ISM dust.  $T_{\text{mw}}$  is set by the diffuse, cold dust component that dominates the total dust mass, while  $T_{\text{peak}}$  is also influenced by the dense, warm dust component in the close vicinity of young star clusters. The former component is typically heated less effectively by young stars than the latter so that  $T_{\text{peak}}$  and  $T_{\text{mw}}$  are not well correlated with each other.

The increase of  $T_{\text{eqv}}$  with redshift is consistent with recent observational evidence, including low number counts of (sub)mm sources in ALMA blind surveys (Bouwens et al. 2016; Casey et al. 2018b, and references therein) and the unusual IRX- $\beta$  relation of high-redshift galaxies (cf. Capak et al. 2015; Ma et al. 2019). However, as we argue in this paper, the rise of  $T_{\text{eqv}}$  with redshift is not simply a sign of dust being hotter at higher redshift, but it reflects the change in SED shape. In particular, higher  $T_{\text{eqv}}$  is often a consequence of a more prominent MIR emission of galaxies at higher redshift, resulting from more active star formation. However, as  $T_{\text{mw}}$  evolves only weakly between  $z = 2$  and 6, the temperature of the majority of the dust component ( $\sim T_{\text{mw}}$ ) does not significantly change despite the change in  $T_{\text{eqv}}$ . In this sense, dust in galaxies with higher  $T_{\text{eqv}}$  is not necessarily *physically* hotter.

In conclusion, dust temperature is important for estimating and probing key physical properties (e.g. dust/gas mass, IR luminosity) and ISM conditions of high-redshift galaxies. A proper interpretation of dust temperatures and their scaling relationships requires taking into account the differences between temperatures derived from the SED shape and the physical, mass-weighted dust temperature. Upcoming facilities, such as JWST, SPICA, and CCAT-prime, will significantly improve our capability of constraining key dust properties of galaxies in the distant Universe.

## ACKNOWLEDGEMENTS

We are grateful to the referee for the careful review and insightful suggestions that have helped improve the quality of this manuscript. We acknowledge the comments from Andreas Faisst, Rob Ivison, Allison Kirkpatrick, Georgios Magdis, and Ian Smail to an early draft. We thank Caitlin Casey for a discussion on Casey et al. (2018a,b) before they were published on arXiv, which had helped the writing of this paper. We are grateful to the technical guidance by Peter Camps and Maarten Baes. LL would like to acknowledge the stimulating atmosphere during the Munich Institute for Astro- and Particle Physics (MIAPP) 2018 programme *The Interstellar Medium of High Redshift Galaxies*, and especially thank the conversation with the programme coordinators. This research was supported by the MIAPP of the Deutsche Forschungsgemeinschaft (DFG) cluster of excellence ‘Origin and Structure of the Universe’. RF acknowledges financial support from the Swiss National Science Foundation (grant no. 157591). Simulations were run with resources provided by the NASA High-End Computing (HEC) Programme

through the NASA Advanced Supercomputing (NAS) Division at Ames Research centre, proposal SMD-14-5492. Additional computing support was provided by HEC allocations SMD-14-5189, SMD-15-5950, by NSF XSEDE allocations AST120025, AST150045, by allocations s697, s698 at the Swiss National Supercomputing Centre (CSCS), and by S3IT resources at the University of Zurich. DK acknowledges support from the NSF grant AST-1715101 and the Cottrell Scholar Award from the Research Corporation for Science Advancement. CAFG was supported by NSF through grants AST-1517491, AST-1715216, and CAREER award AST-1652522; by NASA through grant 17-ATP17-0067; by STScI through grant *HST*-AR-14562.001; and by a Cottrell Scholar Award from the Research Corporation for Science Advancement. PFH was supported by an Alfred P. Sloan Research Fellowship, NASA ATP Grant NNX14AH35G, and NSF Collaborative Research Grant #1411920 and CAREER grant #1455342. EQ was supported in part by a Simons Investigator Award from the Simons Foundation and by NSF grant AST-1715070. The Flatiron Institute is supported by the Simons Foundation.

## REFERENCES

- Aravena M. et al., 2016, *ApJ*, 833, 68  
 Armus L. et al., 2009, *PASP*, 121, 559  
 Azadi M. et al., 2018, *ApJ*, 866, 63  
 Baes M., Camps P., 2015, *Astron. Comput.*, 12, 33  
 Baes M., Verstaappen J., De Looze I., Fritz J., Saftly W., Vidal Pérez E., Stalevski M., Valcke S., 2011, *ApJS*, 196, 22  
 Behrens C., Pallottini A., Ferrara A., Gallerani S., Vallini L., 2018, *MNRAS*, 477, 552  
 Béthermin M. et al., 2015, *A&A*, 573, A113  
 Blain A. W., Smail I., Ivison R. J., Kneib J.-P., 1999, *MNRAS*, 302, 632  
 Blain A. W., Barnard V. E., Chapman S. C., 2003, *MNRAS*, 338, 733  
 Boselli A. et al., 2010, *PASP*, 122, 261  
 Bouwens R. et al., 2016, *ApJ*, 833, 72  
 Broekhoven-Fiene H. et al., 2018, *ApJ*, 852, 73  
 Calzetti D., Kinney A. L., Storchi-Bergmann T., 1994, *ApJ*, 429, 582  
 Calzetti D., Meurer G. R., Bohlin R. C., Garnett D. R., Kinney A. L., Leitherer C., Storchi-Bergmann T., 1997, *AJ*, 114, 1834  
 Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, 533, 682  
 Camps P., Trayford J. W., Baes M., Theuns T., Schaller M., Schaye J., 2016, *MNRAS*, 462, 1057  
 Capak P. L. et al., 2011, *Nature*, 470, 233  
 Capak P. L. et al., 2015, *Nature*, 522, 455  
 Carilli C., Walter F., 2013, *ARA&A*, 51, 105  
 Carlstrom J. E. et al., 2011, *PASP*, 123, 568  
 Carniani S. et al., 2015, *A&A*, 584, A78  
 Casey C. M., 2012, *MNRAS*, 425, 3094  
 Casey C. M., Narayanan D., Cooray A., 2014, *Phys. Rep.*, 541, 45  
 Casey C. M. et al., 2018a, *ApJ*, 862, 77  
 Casey C. M., Hodge J., Zavala J. A., Spilker J., da Cunha E., Stagnu J., Finkelstein S. L., Drew P., 2018b, *ApJ*, 862, 78  
 Cazaux S., Tielens A. G. G. M., 2002, *ApJ*, 575, L29  
 Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, *ApJ*, 622, 772  
 Chary R., Elbaz D., 2001, *ApJ*, 556, 562  
 Chen B., Dai X., Kochanek C. S., Chartas G., 2013, preprint (arXiv:1306.0008)  
 Chen C.-C., Cowie L. L., Barger A. J., Wang W.-H., Williams J. P., 2014, *ApJ*, 789, 12  
 Chiang I.-D., Sandstrom K. M., Chasteney J., Johnson L. C., Leroy A. K., Utomo D., 2018, *ApJ*, 865, 117  
 Cochrane R. K. et al., 2019, *MNRAS*, 488, 1779  
 Conley A. et al., 2011, *ApJ*, 732, L35  
 Conroy C., 2013, *ARA&A*, 51, 393  
 da Cunha E., Charlot S., Elbaz D., 2008, *MNRAS*, 388, 1595

- da Cunha E. et al., 2013, *ApJ*, 766, 13
- da Cunha E. et al., 2015, *ApJ*, 806, 110
- Daddi E., Dannerbauer H., Krips M., Walter F., Dickinson M., Elbaz D., Morrison G. E., 2009, *ApJ*, 695, L176
- Dai X., Kochanek C. S., 2009, *ApJ*, 692, 677
- Dale D. A., Helou G., 2002, *ApJ*, 576, 159
- Dale D. A., Helou G., Contursi A., Silbermann N. A., Kolhatkar S., 2001, *ApJ*, 549, 215
- Dale D. A. et al., 2012, *ApJ*, 745, 95
- Davé R., Finlator K., Oppenheimer B. D., 2011, *MNRAS*, 416, 1354
- De Cia A., Ledoux C., Savaglio S., Schady P., Vreeswijk P. M., 2013, *A&A*, 560, A88
- De Cia A., Ledoux C., Mattsson L., Petitjean P., Srianand R., Gavignaud I., Jenkins E. B., 2016, *A&A*, 596, A97
- De Looze I. et al., 2014, *A&A*, 571, A69
- De Vis P. et al., 2017, *MNRAS*, 471, 1743
- De Vis P. et al., 2019, *A&A*, 623, A5
- Draine B. T., 2006, *ApJ*, 636, 1114
- Draine B. T., Li A., 2007, *ApJ*, 657, 810
- Draine B. T. et al., 2007, *ApJ*, 663, 866
- Dunlop J. S. et al., 2017, *MNRAS*, 466, 861
- Dunne L., Eales S. A., 2001, *MNRAS*, 327, 697
- Dunne L., Eales S., Edmunds M., Ivison R., Alexander P., Clements D. L., 2000, *MNRAS*, 315, 115
- Dwek E., 1998, *ApJ*, 501, 643
- Eales S. et al., 2010, *PASP*, 122, 499
- Egami E. et al., 2018, *Publ. Astron. Soc. Aust.*, 35, 48
- Elbaz D., Cesarsky C. J., Chantal P., Aussel H., Franceschini A., Fadda D., Chary R. R., 2002, *A&A*, 384, 848
- Elbaz D. et al., 2011, *A&A*, 533, A119
- Elbaz D. et al., 2018, *A&A*, 616, A110
- Eldridge J. J., Stanway E. R., 2012, *MNRAS*, 419, 479
- Erb D. K., Shapley A. E., Pettini M., Steidel C. C., Reddy N. A., Adelberger K. L., 2006, *ApJ*, 644, 813
- Faisst A. L. et al., 2016, *ApJ*, 822, 29
- Faisst A. L. et al., 2017, *ApJ*, 847, 21
- Faucher-Giguère C.-A., 2018, *MNRAS*, 473, 3717
- Fazio G. G. et al., 2004, *ApJS*, 154, 10
- Feldmann R., 2017, *MNRAS*, 470, L59
- Feldmann R., Gnedin N. Y., Kravtsov A. V., 2012, *ApJ*, 747, 124
- Feldmann R., Hopkins P. F., Quataert E., Faucher-Giguère C.-A., Kereš D., 2016, *MNRAS*, 458, L14
- Feldmann R., Quataert E., Hopkins P. F., Faucher-Giguère C.-A., Kereš D., 2017, *MNRAS*, 470, 1050
- Ferrara A., Hirashita H., Ouchi M., Fujimoto S., 2017, *MNRAS*, 471, 5018
- Finlator K., Davé R., 2008, *MNRAS*, 385, 2181
- Franco M. et al., 2018, *A&A*, 620, A152
- Fudamoto Y. et al., 2017, *MNRAS*, 472, 483
- Fujimoto S., Ouchi M., Ono Y., Shibuya T., Ishigaki M., Nagai H., Momose R., 2016, *ApJS*, 222, 1
- Galametz M. et al., 2012, *MNRAS*, 425, 763
- Galliano F., Madden S. C., Jones A. P., Wilson C. D., Bernard J.-P., 2005, *A&A*, 434, 867
- Galliano F. et al., 2011, *A&A*, 536, A88 (G11)
- Geach J. E. et al., 2013, *MNRAS*, 432, 53
- Geach J. E. et al., 2017, *MNRAS*, 465, 1789
- Gill S. P. D., Knebe A., Gibson B. K., 2004, *MNRAS*, 351, 399
- Gonzalez-Alfonso E., Smith H. A., Fischer J., Cernicharo J., 2004, *ApJ*, 613, 247
- Griffin M. J. et al., 2010, *A&A*, 518, L3
- Grogan N. A. et al., 2011, *ApJS*, 197, 35
- Groves B., Dopita M. A., Sutherland R. S., Kewley L. J., Fischera J., Leitherer C., Brandl B., van Breugel W., 2008, *ApJS*, 176, 438
- Groves B. A. et al., 2015, *ApJ*, 799, 96
- Hahn O., Abel T., 2011, *MNRAS*, 415, 2101
- Harrington K. C. et al., 2018, *MNRAS*, 474, 3866
- Harvey P. M. et al., 2013, *ApJ*, 764, 133
- Hashimoto T. et al., 2018, *Nature*, 557, 392
- Hatsukade B., Ohta K., Seko A., Yabe K., Akiyama M., 2013, *ApJ*, 769, L27
- Hatsukade B. et al., 2016, *PASJ*, 68, 36
- Hatsukade B. et al., 2018, *PASJ*, 70, 105
- Hayward C. C., Smith D. J. B., 2015, *MNRAS*, 446, 1512
- Hayward C. C., Kereš D., Jonsson P., Narayanan D., Cox T. J., Hernquist L., 2011, *ApJ*, 743, 159
- Hayward C. C., Jonsson P., Kereš D., Magnelli B., Hernquist L., Cox T. J., 2012, *MNRAS*, 424, 951
- Heckman T. M., Robert C., Leitherer C., Garnett D. R., van der Rydt F., 1998, *ApJ*, 503, 646
- Herrera-Camus R. et al., 2012, *ApJ*, 752, 112
- Hildebrand R. H., 1983, *QJRAS*, 24, 267
- Hinshaw G. et al., 2013, *ApJS*, 208, 19
- Hirashita H., Nozawa T., Villaume A., Srinivasan S., 2015, *MNRAS*, 454, 1620
- Hodge J. A. et al., 2013, *ApJ*, 768, 91
- Holland W. S. et al., 1999, *MNRAS*, 303, 659
- Holland W. S. et al., 2013, *MNRAS*, 430, 2513
- Hopkins P. F., 2013, *MNRAS*, 428, 2840
- Hopkins P. F., 2015, *MNRAS*, 450, 53
- Hopkins P. F., Quataert E., Murray N., 2012, *MNRAS*, 421, 3522
- Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, 445, 581
- Hughes T. M. et al., 2017, *MNRAS*, 468, L103
- Hwang H. S. et al., 2010, *MNRAS*, 409, 75
- Inoue A. K., 2003, *PASJ*, 55, 901
- Issa M. R., MacLaren I., Wolfendale A. W., 1990, *A&A*, 236, 237
- Ivison R. J. et al., 2010, *A&A*, 518, L35
- Ivison R. J. et al., 2016, *ApJ*, 832, 78
- Izumi T. et al., 2019, preprint ([arXiv:1904.07345](https://arxiv.org/abs/1904.07345))
- James A., Dunne L., Eales S., Edmunds M. G., 2002, *MNRAS*, 335, 753
- Janowiecki S., Cortese L., Catinella B., Goodwin A. J., 2018, *MNRAS*, 476, 1390
- Jonsson P., Groves B. A., Cox T. J., 2010, *MNRAS*, 403, 17
- Kaasinen M. et al., 2019, *ApJ*, 880, 15
- Karim A. et al., 2013, *MNRAS*, 432, 2
- Katz H., Laporte N., Ellis R. S., Devriendt J., Slyz A., 2019, *MNRAS*, 484, 4054
- Kessler M. F. et al., 1996, *A&A*, 315, L27
- Kewley L. J., Dopita M. A., 2002, *ApJS*, 142, 35
- Kewley L. J., Ellison S. L., 2008, *ApJ*, 681, 1183
- Kinney A. L., Bohlin R. C., Calzetti D., Panagia N., Wyse R. F. G., 1993, *ApJS*, 86, 5
- Kirkpatrick A. et al., 2012, *ApJ*, 759, 139
- Kirkpatrick A., Pope A., Sajina A., Roebuck E., Yan L., Armus L., Díaz-Santos T., Stierwalt S., 2015, *ApJ*, 814, 9
- Kirkpatrick A. et al., 2017, *ApJ*, 843, 71
- Knollmann S. R., Knebe A., 2009, *ApJS*, 182, 608
- Koekemoer A. M. et al., 2011, *ApJS*, 197, 36
- Kovács A. et al., 2010, *ApJ*, 717, 29
- Koyama Y. et al., 2015, *MNRAS*, 453, 879
- Kriek M., Conroy C., 2013, *ApJ*, 775, L16
- Lambrides E. L., Petric A. O., Tchernyshyov K., Zakamska N. L., Watts D. J., 2019, *MNRAS*, 487, 1823
- Laporte N. et al., 2017, *ApJ*, 837, L21
- Lee N. et al., 2013, *ApJ*, 778, 131
- Liang L., Feldmann R., Faucher-Giguère C.-A., Kereš D., Hopkins P. F., Hayward C. C., Quataert E., Scoville N. Z., 2018, *MNRAS*, 478, L83
- Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, *ApJ*, 772, 119
- Lisenfeld U., Ferrara A., 1998, *ApJ*, 496, 145
- Lombardi M., Bouy H., Alves J., Lada C. J., 2014, *A&A*, 566, A45
- Lutz D., 2014, *ARA&A*, 52, 373
- Lutz D. et al., 2011, *A&A*, 532, A90
- Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, *MNRAS*, 456, 2140
- Ma X. et al., 2019, *MNRAS*, 487, 1844



- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Magdis G. E. et al., 2010, *MNRAS*, 409, 22
- Magdis G. E. et al., 2012, *ApJ*, 760, 6
- Magdis G. E. et al., 2013, *A&A*, 558, A136
- Magdis G. E. et al., 2017, *A&A*, 603, A93
- Magnelli B. et al., 2012, *A&A*, 539, A155
- Magnelli B. et al., 2013, *A&A*, 553, A132
- Magnelli B. et al., 2014, *A&A*, 561, A86
- Maiolino R. et al., 2008, *A&A*, 488, 463
- Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, *MNRAS*, 408, 2115
- Mathis J. S., Mezger P. G., Panagia N., 1983, *A&A*, 128, 212
- Matthee J. et al., 2019, preprint ([arXiv:1903.08171](https://arxiv.org/abs/1903.08171))
- Mattsson L., De Cia A., Andersen A. C., Zafar T., 2014, *MNRAS*, 440, 1562
- McAlpine S. et al., 2019, *MNRAS*, 488, 2440
- McKee C. F., Ostriker E. C., 2007, *ARA&A*, 45, 565
- McKinnon R., Torrey P., Vogelsberger M., 2016, *MNRAS*, 457, 3775
- Meurer G. R., Heckman T. M., Leitherer C., Kinney A., Robert C., Garnett D. R., 1995, *AJ*, 110, 2665
- Meurer G. R., Heckman T. M., Calzetti D., 1999, *ApJ*, 521, 64
- Miller T. B. et al., 2018, *Nature*, 556, 469
- Murata K. et al., 2014, *A&A*, 566, A136
- Murray N., Quataert E., Thompson T. A., 2005, *ApJ*, 618, 569
- Narayanan D., Hayward C. C., Cox T. J., Hernquist L., Jonsson P., Younger J. D., Groves B., 2010, *MNRAS*, 401, 1613
- Narayanan D. et al., 2015, *Nature*, 525, 496
- Narayanan D., Davé R., Johnson B. D., Thompson R., Conroy C., Geach J., 2018a, *MNRAS*, 474, 1718
- Narayanan D., Conroy C., Davé R., Johnson B. D., Popping G., 2018b, *ApJ*, 869, 70
- Nguyen H. T. et al., 2010, *A&A*, 518, L5
- Nordon R. et al., 2012, *ApJ*, 745, 182
- Oliver S. J. et al., 2012, *MNRAS*, 424, 1614
- Ono Y., Ouchi M., Kurono Y., Momose R., 2014, *ApJ*, 795, 5
- Onodera M. et al., 2016, *ApJ*, 822, 42
- Oteo I. et al., 2017, preprint ([arXiv:1709.04191](https://arxiv.org/abs/1709.04191))
- Oteo I. et al., 2018, *ApJ*, 856, 72
- Patil P., Nyland K., Lacy M., Farrah D., Afonso J., Barkhouse W., Surace J., 2019, *ApJ*, 871, 109
- Pavesi R. et al., 2018, *ApJ*, 861, 43
- Pilbratt G. L. et al., 2010, *A&A*, 518, L1
- Poglitsch A. et al., 2010, *A&A*, 518, L2
- Pope A. et al., 2008, *ApJ*, 675, 1171
- Popping G., Puglisi A., Norman C. A., 2017, *MNRAS*, 472, 2315
- Price S. H. et al., 2014, *ApJ*, 788, 86
- Privon G. C., Narayanan D., Davé R., 2018, *ApJ*, 867, 102
- Reddy N. A., Steidel C. C., Fadda D., Yan L., Pettini M., Shapley A. E., Erb D. K., Adelberger K. L., 2006, *ApJ*, 644, 792
- Reddy N. A., Erb D. K., Pettini M., Steidel C. C., Shapley A. E., 2010, *ApJ*, 712, 1070
- Rémy-Ruyer A. et al., 2014, *A&A*, 563, A31
- Rennehan D., Babul A., Hayward C. C., Bottrell C., Hani M. H., Chapman S. C., 2019, preprint ([arXiv:1907.00977](https://arxiv.org/abs/1907.00977))
- Riechers D. A. et al., 2013, *Nature*, 496, 329
- Rieke G. H., Alonso-Herrero A., Weiner B. J., Pérez-González P. G., Blaylock M., Donley J. L., Marcillac D., 2009, *ApJ*, 692, 556
- Rigopoulou D., Pereira-Santaella M., Magdis G. E., Cooray A., Farrah D., Marques-Chaves R., Perez-Fournon I., Riechers D., 2018, *MNRAS*, 473, 20
- Riguccini L. et al., 2015, *MNRAS*, 452, 470
- Roebuck E., Sajina A., Hayward C. C., Pope A., Kirkpatrick A., Hernquist L., Yan L., 2016, *ApJ*, 833, 60
- Rujopakarn W., Rieke G. H., Weiner B. J., Pérez-González P., Rex M., Walth G. L., Kartaltepe J. S., 2013, *ApJ*, 767, 73
- Safarzadeh M., Hayward C. C., Ferguson H. C., Somerville R. S., 2016, *ApJ*, 818, 62
- Santini P. et al., 2009, *A&A*, 504, 751
- Schinnerer E. et al., 2016, *ApJ*, 833, 112
- Schreiber C. et al., 2015, *A&A*, 575, A74
- Schreiber C., Elbaz D., Pannella M., Ciesla L., Wang T., Franco M., 2018, *A&A*, 609, A30
- Scoville N. et al., 2014, *ApJ*, 783, 84
- Scoville N. et al., 2016, *ApJ*, 820, 83
- Scoville N. et al., 2017a, *ApJ*, 836, 66
- Scoville N. et al., 2017b, *ApJ*, 837, 150
- Scoville N. Z., 2013, in Falcón-Barroso J., Knapen J. H., eds, *Secular Evolution of Galaxies*. Cambridge Univ. Press, Cambridge, UK, p. 491
- Scoville N. Z., Kwan J., 1976, *ApJ*, 206, 718
- Shetty R., Glover S. C., Dullemond C. P., Klessen R. S., 2011, *MNRAS*, 412, 1686
- Shipley H. V., Papovich C., Rieke G. H., Brown M. J. I., Moustakas J., 2016, *ApJ*, 818, 60
- Siana B. et al., 2009, *ApJ*, 698, 1273
- Siebenmorgen R., Krügel E., 2007, *A&A*, 461, 445
- Siebenmorgen R., Krügel E., Spoon H. W. W., 2004, *A&A*, 414, 123
- Simpson J. M. et al., 2017, *ApJ*, 839, 58 (S17)
- Smail I., Ivison R. J., Blain A. W., 1997, *ApJ*, 490, L5
- Sobral D., Smail I., Best P. N., Geach J. E., Matsuda Y., Stott J. P., Cirasuolo M., Kurk J., 2013, *MNRAS*, 428, 1128
- Sparre M., Hayward C. C., Feldmann R., Faucher-Giguère C.-A., Muratov A. L., Kereš D., Hopkins P. F., 2017, *MNRAS*, 466, 88
- Spinoglio L. et al., 2017, *Publ. Astron. Soc. Aust.*, 34, e057
- Steidel C. C. et al., 2014, *ApJ*, 795, 165
- Stierwalt S. et al., 2014, *ApJ*, 790, 122
- Strandet M. L. et al., 2016, *ApJ*, 822, 80
- Swinbank A. M. et al., 2014, *MNRAS*, 438, 1267
- Symeonidis M. et al., 2013, *MNRAS*, 431, 2317 (S13)
- Takeuchi T. T., Yuan F.-T., Ikegama A., Murata K. L., Inoue A. K., 2012, *ApJ*, 755, 144
- Thomson A. P. et al., 2017, *ApJ*, 838, 119 (T17)
- Trayford J. W. et al., 2017, *MNRAS*, 470, 771
- Tremonti C. A. et al., 2004, *ApJ*, 613, 898
- Tuffs R. J., Popescu C. C., Völk H. J., Kylafis N. D., Dopita M. A., 2004, *A&A*, 419, 821
- Utomo D., Chiang I.-D., Leroy A. K., Sandstrom K. M., Chastenot J., 2019, *ApJ*, 874, 141
- Venemans B. P. et al., 2017, *ApJ*, 851, L8
- Walcher J., Groves B., Budavári T., Dale D., 2011, *Ap&SS*, 331, 1
- Walter F. et al., 2012, *Nature*, 486, 233
- Walter F. et al., 2016, *ApJ*, 833, 67
- Watson D., 2011, *A&A*, 533, A16
- Watson D., Christensen L., Knudsen K. K., Richard J., Gallazzi A., Michałowski M. J., 2015, *Nature*, 519, 327
- Weingartner J. C., Draine B. T., 2001, *ApJ*, 548, 296
- Weiß A. et al., 2013, *ApJ*, 767, 88
- Wiklund T. et al., 2019, *ApJ*, 878, 83
- Wild V., Charlot S., Brinchmann J., Heckman T., Vince O., Pacifici C., Chevallard J., 2011, *MNRAS*, 417, 1760
- Wilkins S. M., Gonzalez-Perez V., Lacey C. G., Baugh C. M., 2012, *MNRAS*, 427, 1490
- Wilson G. W. et al., 2008, *MNRAS*, 386, 807
- Wiseman P., Schady P., Bolmer J., Krühler T., Yates R. M., Greiner J., Fynbo J. P. U., 2017, *A&A*, 599, A24
- Zafar T., Watson D., 2013, *A&A*, 560, A26
- Zahid H. J., Kewley L. J., Bresolin F., 2011, *ApJ*, 730, 137
- Zavala J. A. et al., 2018a, *Nat. Astron.*, 2, 56
- Zavala J. A., et al., 2018b, *MNRAS*, 475, 5585 (Z18)
- Zavala J. A., Casey C. M., Cunha E. d., Spilker J., Staguhn J., Hodge J., Drew P. M., 2018c, *ApJ*, 869, 71

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.